Interactions Between the Adeno- Associated Virus Replication Proteins and Host Cellular Factors

Smith, Sarah Claire

Awarding institution:
King's College London

Interactions Between the Adeno-Associated Virus Replication Proteins and Host Cellular Factors

Sarah C Smith

This thesis is submitted for the degree of Doctor of Philosophy
at King's College London

Department of Infectious Diseases
Faculty of Life Sciences & Medicine

- September 2016 -

To my father Peter and my grandfather Louis

ABSTRACT

Adeno-associated virus (AAV) is a small, ssDNA virus with a unique biphasic life cycle in which productive replication is dependent on both helper virus and host cellular factors. Unable to replicate autonomously, infection by AAV alone leads to the establishment of latency through either integration of the viral genome, or long-term episomal persistence. The complex AAV life cycle is orchestrated almost entirely by four isoforms of a single multifunctional viral nonstructural replication (Rep) protein. While it is known that Rep must interact with a multitude of host factors in order to complete the AAV life cycle, little is yet known about the nature of such interactions with respect to AAV gene regulation and the establishment of latency. Here, we used a screening method called BioID to identify new interaction partners for the Rep proteins, resulting in the identification of a number of interesting candidates involved in gene regulation, including the transcriptional corepressor KAP1. We show that KAP1 binds the latent AAV2 genome at the *rep* ORF, leading to trimethylation of AAV2-associated H3K9. We present evidence that helper viruses target KAP1 for degradation and demonstrate that antagonism of PP1α by Rep52/Rep78 further counteracts KAP1-mediated repression by enhancing nuclear levels of phosphorylated KAP1-S824, and that this interaction is essential for AAV2 transcription and replication. This work challenges the currently held model for AAV latency, and introduces not only a new viral mechanism for the counteraction of KAP1 repression, but also the notion that KAP1 targeting represents a conserved requirement for replication among DNA viruses.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Els Henckaerts. Your constant support, seemingly endless patience, gentle but firm guidance, and belief in my abilities has truly given me the opportunity to learn to stand on my own two feet while also allowing me the space to fall without taking anyone else down with me. I have learned so much from you, scientifically and otherwise. To my second supervisor, Michael Linden, discussing science with you has been an illuminating experience, both in the way that it leads each and every time to the discovery of a new concept for me, and in how it has changed the way that I approach and assess problems on my own.

I am also particularly thankful to my thesis committee, Stuart Neil, Chad Swanson, and Fiona Watt, who challenged me and ensured that I never strayed too far from the right track. In particular Stuart and Chad, whose input in both technical and scientific aspects of the project has been invaluable.

A special thanks also to Helen Rowe, without whom none of the work on KAP1 would have been possible. Her material and scientific contributions to the project have been immense, and I am deeply grateful.

I would also like to extend my deepest gratitude to my fantastic lab-mates over these past four years. Martino, I don't know where I would have ended up without your help! From discussing the nuances of our nemesis Rep, to trouble-shooting experiments or just making me smile with silly cat videos, you have been an essential part of this process for me. To Julie, whose great friendship has helped me immeasurably through the more personal struggles of a PhD, and whose enthusiastic "Bonjouuuur!" every morning was enough to lift even the lowest of spirits. To André, whose laid-back approach to life taught me that sometimes, it's okay to slow down. To Leti, who taught me everything I know about bacterial culture and molecular cloning, and also a thing or two regarding the finer points of Spanish profanity. And finally, to Núria, whose smile and witty remarks made every day fun.

On a more personal level, I would also like to thank my partner Duncan. You have seen me begin and now end this process, and your unconditional support throughout has been truly touching. Your confidence in me and in the significance of this endeavor often ran deeper than my own, and it gave me the strength I needed when I did not have it.

And finally, to my family, thank you. Thank you for so much. Thank you for your forgiveness when I tested the very boundaries of family ties. Thank you for giving me the room to be myself, and for loving me in spite of it. Thank you for being the most generous people on Earth, in every way. To my father in particular, thank you for introducing me to the wonders of science, mathematics, and the cosmos at the tender age of five through your peerless bedtime stories – I still remember Peg-Leg, his mossy knee, and my cube roots! You gave me the gift of fascination, and without this, I would be nowhere.

ABBREVIATONS

AAA ATPase	ATPase of the triple A helicase family
AAP	Assembly activating protein
AAVS1	AAV integration site 1
AAV	Adeno-associated virus
AAVR	AAV receptor
Ad	Adenovirus
AdMLP	Ad major late transcription promoter
ANP32A	Acidic nuclear protein 32A
ANP32B	Acidic nuclear protein 32B
ATM	Ataxia telangiectasia mutated kinase
ATP	Adenosine triphosphate
ATR	ATM and Rad3-related kinase
BioID	Proximity-dependent biotin identification
bp	Base pairs
BR	Basic regions
BSA	Bovine serum albumin
BUdR	Bromodeoxyuridine
CHD3	Chromodomain-helicase-DNA-binding protein 3
C-terminus	Carboxy-terminus
ChIP	Chromatin immunoprecipitation
CHK1	Checkpoint kinase 1
CHK2	Checkpoint kinase 2
CIP	Calf intestinal phosphatase
CL-IP	Cross-linked immunoprecipitation
CMV	Cytomegalovirus
Co-IP	Co-immunoprecipitation
CPE	Cytopathic effect
CPSF6	Cleavage and polyadenylation specific factor 6
CREB	cAMP response element-binding protein
CRE	cAMP response element
CTD	Carboxy terminal domain
DBP	DNA binding protein of adenovirus (product of E2A gene)
DC	Dendritic cell
DDR	DNA damage response
DMEM	Dulbecco’s modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNA-PK	DNA-dependent protein kinase

dNTP	Deoxynucleotide triphosphate
DSB	Double strand breaks
dsDNA	Double-stranded DNA
EBV	Epstein-Barr virus
EC	Embryonic carcinoma
EDTA	1-(4-Aminobenzyl)ethylenediamine-N,N,N'N'-tetraacetic acid
EGFR	Epidermal growth factor receptor
ERV	Endogenous retroviruses
ES	Embryonic stem
FBS	Fetal bovine serum
FGFR-1	Fibroblast growth receptor 1
GFP	Green fluorescent protein
GLIC/GLEEC	Clathrin-Independent Carriers/GPI-Enriched Endocytic
GMP	Good manufacturing practice
hFIX	Human factor IX
H	Hour
HDAC	Histone deacetylase
HEK293T	Human embryonic kidney 293T cells
HeLa	Henrietta Lacks (Human epitheloid carcinoma cells)
HIV	Human immunodeficiency virus
HMG1	High mobility group protein 1
HP1BD	HP1 binding domain
HP1	Heterochromatin protein 1
HPV	Human papillomavirus
HR	Homologous recombination
HRP	Horseradish peroxidase
HSC	Hematopoietic stem cell
HSP90	Heat shock protein 90
HSPG	Heparan sulfate proteoglycans
HSV-1	Herpes simplex virus type 1
H3K9me3	Trimethylated histone 3 lysine 9
IE	Immediate early
IN	Integrase
ITR	Inverted terminal repeat
IU	Infectious unit
kB	Kilobase
kDa	Kilodalton
KRAB-ZFP	Krupple-associated Box-domain containing ZNF protein
KRV	Kilham's rat virus
KSHV	Kaposi's sarcoma-associated herpesvirus

LaA	Lamin-A
LANA	Latency associated nuclear antigen
LB	Luria broth
LCA2	Leber Congenital Amaurosis 2
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LTag	Large T antigen of SV40
LTR	Long terminal repeat
MBS85	Myosin-binding subunit 5
MCM	Minichromosome maintenance complex
MDC1	Mediator of damage checkpoint 1
MDM2	Mouse double minute 2 homolog
mg	Milligram
ml	Milliliter
MLV	Murine leukemia virus
MLTF	Major late transcription factor
mM	Millimolar
MOI	Multiplicity of infection
mRNA	Messenger RNA
MRN	Mre11/Rad50/Nbs1 complex
MVM	Minute virus of mice
nAb	Neutralizing antibodies
NE	Nuclear envelope
NEB	New England Biolabs
ng	Nanogram
NELF	Negative elongation factor
NHEJ	Non-homologous end joining
NLS	Nuclear localization signal
nM	Nanomolar
NoLS	Nucleolar localization signal
NPC	Nuclear pore complex
NTP	Nucleotide triphosphate
NuRD	Nucleosome remodeling and histone deacetylation
OBD	Origin binding domain
ORF	Open reading frame
Ori	Origin of replication
PB	Plant homeodomain and bromodomain
PBS	Phosphate buffered saline
PCNA	Proliferating nuclear antigen
PCR	Polymerase chain reaction
PDGFR	Platelet derived growth factor receptor

PEI	Polyethylenimine
PFA	Paraformaldehyde
pH	Power of hydrogen
PIC	Pre-initiation complex
PFU	Plaque forming unit
PHD	Plant homeodomain
PKI	Protein kinase inhibitor
PLA2	Phospholipase 2
POLD	DNA polymerase delta
PP1	Protein phosphatase 1
PPP1R12C	Protein phosphatase 1 regulatory subunit 12C
PRG	Primary response gene
PRR	Pattern recognition receptor
qPCR	Quantitative polymerase chain reaction
rAAV	Recombinant AAV vector
RBS	Rep binding site
RC	Replication centers
RCR	Rolling circle replication
RDBP	RNA binding protein RD (NELF-E)
RE	Restriction enzyme
RFC	Replication factor C
RING	Really interesting new gene
RNA	Ribonucleic acid
RNA pol II	RNA polymerase II
RPE65	Retinal pigment epithelium-specific 65-kDa
RT	Reverse transcriptase
RUVBL1	RuvB-like 1
scAAV	self-complementary rAAV
SDS-PAGE	Sodium dodecyl sulfate polyacrylamaide gel ectrophoresis
SEM	Standard error of the mean
SETDB1	SET domain, bifurcated 1
SF1	Splicing factor 1
SF3	Superfamily 3
SIM	SUMO interacting motif
siRNA	Small interfering RNA
shRNA	Small hairpin RNA
snRNP	Small nuclear ribonucleoprotein
SNW1	SNW domain containing 1
ssDNA	single-stranded DNA
SUMO	Small ubiquitin-like modifier

SV40	Simian virus 40
TAF/Set	Template activating factor I/Set oncoprotein
Tat	Trans-activator of transcription
TBP	TATA binding protein
TCERG1	Transcriptional elongation regulator 1
TE	Transposable elements
TF	Transcription factor
TGN	Trans-golgi network
TIF1	Transcription initiation factor 1
TRIM	Tripartite motif
TRS	Terminal resolution site
TSS	Transcription start site
URR	Upstream regulatory region
UTR	Untranslated region
VP	Viral protein
VSV-G	Vesicular stomatitis virus G glycoprotein
w/v	weight/volume
WT	Wild-type
YY1	Ying Yang 1
Y2H	Yeast two hybrid
ZNF	Zinc finger
α	Anti
µg	Microgram
µl	Microliter
µM	Micromolar

INDEX OF FIGURES

INDEX OF TABLES

CHAPTER 1: INTRODUCTION

1.1 ADENO-ASSOCIATED VIRUS

1.1.1 A History of AAV Discovery

Adeno-associated virus (AAV) was originally discovered 50 years ago as a contaminant of both human and simian adenovirus preparations (Atchison et al., 1965; Hoggan et al., 1966). It was initially debated whether the observed hexagonal particles approximately 24 nm in diameter were representative of internal components from mature adenoviral particles shed *in vitro* (Mayor et al., 1965), or an independent virus. The latter hypothesis was substantiated through studies into the biological and immunological properties of these particles, which demonstrated that AAV was distinct from adenovirus in thermal stability and serological profile. These studies also revealed the defective nature of AAV by demonstrating that infection with AAV alone elicited no cytopathic effect, and furthermore that it was dependent on heterologous adenovirus for replication (Atchison et al., 1965; Hoggan et al., 1966). The observation that infection with adenovirus of African green monkey kidney cells as well as certain primary cultures of human embryonic kidney cells resulted in the rescue of infectious AAV later led to the discovery that, in the absence of adenovirus infection, AAV establishes a persistent latent infection through integration (Cheung et al., 1980; Hoggan et al., 1972).

Initial studies into the composition of the AAV genome yielded conflicting results. Biochemical and hydrodynamic analysis of extracted DNA was consistent with a double-stranded configuration with a molecular weight of 3×10^6 bases (Rose et al., 1966), while acridine orange staining of purified virions suggested a single-stranded secondary structure (Mayor and Melnick, 1966). A detailed comparison of the DNA composition and packaging capacity of AAV with the Minute Virus of Mice (MVM) led to the suggestion that AAV genomes were individually packaged as positive and negative single-stranded DNA molecules that annealed to form duplex structures upon extraction (Crawford et al., 1969). Elegant experiments utilizing the density

labeling of viral DNA with bromodeoxyuridine (BUdR) revealed that a mixture of BUdR-substituted (heavy) and unsubstituted (light) extracted DNA resulted in the formation of DNA density hybrids, strongly supporting this hypothesis (Rose et al., 1969). Confirmation was provided upon the subsequent isolation of the individual genomes, representing the first documented instance of the separation of complementary strands of DNA of an animal virus (Berns and Rose, 1970). Base compositional analysis of isolated genomes then revealed that only the negative sense DNA strand served as a template for transcription (Carter et al., 1972; Rose and Koczot, 1971).

Building upon this work, the observation that isolated single-stranded AAV genomes eluted from hydroxylapatite as if double-stranded led to the deduction that a small region of the genome was partially self complementary (Figure 2, pg. 27) (Carter et al., 1972). This was confirmed when beautiful electron micrographs of single-stranded AAV DNA revealed the presence of various linear and circular DNA concatamers, consistent with the self-annealing of complementary sequences at or near the ends of plus or minus strands of AAV DNA (Koczot et al., 1973). These regions of complementary DNA were later shown by restriction enzyme (RE) digestion to exist both in a linear conformation and as a cruciform structure (Carter et al., 1975; Denhardt et al., 1976), and more detailed studies utilizing RE digest combined with [^{3}H]thymidine pulse-labeling of AAV DNA demonstrated that these regions served as the origin for viral replication (Ori) through a unique mechanism in which DNA synthesis was initiated by the self-priming of these terminal sequences (Hauswirth and Berns, 1978; Straus et al., 1976). Sequencing of these regions, subsequently termed inverted terminal repeats (ITR), determined that they were 145 nucleotides in length, and that the first 125 nucleotides were palindromic (Lusby et al., 1980).

Work by Cathy Laughlin (Laughlin et al., 1979b) and Carol Marcus-Sekura (Marcus et al., 1981) was instrumental in the discovery of the basic transcriptional map for AAV. Together, these studies established the existence of three overlapping mRNA families transcribed from promoters at map units 5, 19, and 40 (p5, p19, and p40), which existed in both spliced and unspliced forms. Furthermore, these studies showed that the smaller 2.3 kb mRNA transcribed from p40 was translated to yield the structural capsid

proteins, VP1, VP2, and VP3 (Jay et al., 1981; Lubeck et al., 1979), which had been described several years previously (Johnson et al., 1972; Rose et al., 1971). It was still unknown at this time what gene products resulted from the p5- and p19-driven transcripts. The existence of a virally encoded protein necessary for replication was made apparent through studies into the replication dynamics of fully infectious AAV particles and naturally occurring defective interfering (DI) AAV particles, which contain aberrant genomes in which an essential gene is deleted and are therefore inherently non-infectious. The observation that AAV DI particle replication could be rescued in the presence of fully functional AAV led the field to suspect the existence of a nonstructural AAV replication protein, missing in the DI genome (Laughlin et al., 1979a). This was confirmed only several years later however, after the cloning and sequencing of the AAV genome.

The nonpathogenic nature of AAV and its ability to establish a persistent latent infection in the absence of adenovirus began to draw attention to its potential for development as a vector for gene therapy. The construction of molecular clones in 1982 (Laughlin et al., 1983; Samulski et al., 1982) was crucial to the advent of this movement as it allowed for the detailed genetic analysis of the viral genome. Transfection of plasmids encoding the full AAV genome into cells infected with adenovirus led to the rescue and replication of infectious AAV without the need for prior excision of the viral genomes. This demonstrated not only the efficiency of the plasmid replication system, but also the fortuitous result that AAV plasmids represented a model for the rescue of integrated latent viral genomes (Laughlin et al., 1983; Samulski et al., 1982). The observation that clones derived from a GC-tailing approach (Samulski et al., 1982) – the majority of which lacked intact ITRs due to recombination between the GC tails and GC rich regions within the ITR – also gave rise to infectious clones led to the characterization of a mechanism for gene recovery by AAV in which duplex plasmids containing at least one intact ITR provide the template for correction of the defective ITR during replication (Samulski et al., 1983). Samulski et al. further made the prescient prediction that the activity of a cleavage enzyme would be necessary to recognize and resolve the Holiday structures predicted to result from this mechanism (Samulski et al., 1983). It

was not however until sequencing of the full AAV genome (Srivastava et al., 1983) rendered the analysis of mutant replication phenotypes possible that the need for at least one nonstructural AAV protein for viral DNA replication was clearly demonstrated (Hermonat et al., 1984; Tratschin et al., 1984). The Rep proteins were later identified (Mendelson et al., 1986), and their purification (Im and Muzyczka, 1992) led to the characterization of the remarkable interaction between the Rep proteins and the viral ITRs.

These mutagenesis studies also paved the way for the development of AAV vectors by demonstrating that the ITRs were the only *cis*-acting function necessary for replication (Hermonat et al., 1984; Tratschin et al., 1984). Both Rep and Cap function could be provided *in trans*, meaning that the entire viral genome could be replaced with a transgene of interest. The first recombinant AAV vectors (rAAV) consisted of plasmids in which only the capsid gene was replaced with a transgene, meaning the vectors generated had a rep+cap- phenotype. Plasmids providing the complementing Rep and Cap functions consisted either of the entire wild-type viral sequence, or had considerable homology with the AAV vectors. These were transfected in the presence of adenovirus to generate rAAV particles, and as a result, all early rAAV preparations were contaminated with varying levels of wild type AAV (wtAAV) as well as heat-inactivated adenovirus. Nevertheless, these early vectors provided a proof-of-concept that rAAV could be generated and used to introduce and express foreign transgenes in mammalian cells (Hermonat et al., 1984; Tratschin et al., 1984). Further advancements in vector production eventually led the generation of rAAV preparations that were free from contamination with either wtAAV or adenovirus, leading to the first *in vivo* studies in 1993 (Flotte et al., 1993). This was followed by the first clinical trial in patients with cystic fibrosis 10 years later (Flotte et al., 2003). Further developments in vector design over the past 20 years have been aimed to increase packaging capacity, transduction efficiency, and tissue specificity. As a result, AAV gene therapy today is demonstrating great success in multiple ongoing clinical trials, highlighting the enormous potential for this therapy in the treatment of debilitating monogenic diseases.

Another aspect of AAV biology that was of intense interest was the viral mechanism for latent persistence. It had previously been shown in

Detroit 6 cells that infection with AAV in the absence of adenovirus led to the establishment of latency through the integration of the viral genome in head-to-tail concatemer formation (Cheung et al., 1980; Hoggan et al., 1972; Laughlin et al., 1986). Similar results were observed in Detroit 6 and KB cells that had been transduced with a Rep-expressing neomycin AAV vector (McLaughlin et al., 1988), however it was noted that these integration events could also have resulted from wtAAV contamination. When sequences flanking the viral-cellular junctions of wtAAV integrants in Detroit 6 cells were used to screen various cell types containing integrated wtAAV or rAAV DNA, it was discovered that most integration events had occurred at a defined region of human chromosome 19, subsequently termed *AAVS1* (Kotin et al., 1990). We now know however that integration of Rep-deficient rAAV genomes occurs only infrequently and without specificity (Kearns et al., 1996), as integration is dependent on the origin binding and endonuclease activities of the large Rep proteins (Im and Muzyczka, 1990a; Linden et al., 1996b; Weitzman et al., 1994), and that the persistence of rAAV genomes occurs mostly in the form of episomal head-to-tail concatemers (Kearns et al., 1996; Schnepp et al., 2003). The advent of additional techniques has since led to the detailed molecular characterization of wtAAV integration into *AAVS1* (Giraud et al., 1994; Henckaerts et al., 2009; Kotin et al., 1990, 1991; Linden et al., 1996b).

The wealth of knowledge with respect to AAV biology has grown exponentially with the emerging development of AAV vector gene therapy. As a result however, the focus has been mostly on improving vector design, and not necessarily on the dissection of basic AAV biology. Various challenges remain in bringing AAV vectors more widely into the clinics. In particular, high doses are currently necessary to effectively target large human organ systems. This presents a real challenge at the level of vector production, and also increases the ordinarily low risk of complications related to vector immunogenicity. As a result, there is a strong drive to establish methods both for the scaling-up of vector production, as well as for the design of vectors with greater bioactivity. Significant enhancements in AAV vector design have previously been derived from an enhanced understanding of basic AAV biology, however there are many aspects of AAV biology that

still remain elusive. Of particular relevance to AAV vectors, whose biology may mimic that of latent viral genomes, little is known about the mechanisms involved in the regulation of, and release from, AAV latency. Understanding the changes that occur at the level of viral chromatin and the protein interactions involved will be instrumental not only in laying the foundation for improved vector design, but also in shedding light on the fascinating interplay between AAV, its helper virus, and a host with which it appears to have so perfectly coevolved.

1.1.2 AAV – the virus

1.1.2.1 The Taxonomy

AAV belongs to a family of small, icosahedral, and non-enveloped DNA viruses called Parvoviridae, whose common characteristics include linear ssDNA genomes terminating in hairpins that render the genome self-priming. The history of this family began in 1959 with the isolation of Kilham rat virus (KRV) from lysates of experimental rat tumors (Kilham and Olivier, 1959). Over the next 10 years, various other similar particles were isolated as contaminants of laboratory tumors, cell cultures, and viral stocks (Atchison et al., 1965; Hoggan et al., 1966; Kilham and Olivier, 1959; Toolan, 1961), each causing variable levels of pathology when injected into rodents. Isolation of another physically similar virus from the larvae of the greater wax moth led to the discovery that these viruses might not be confined to vertebrate hosts. Based on this, the Parvoviridae family was first divided into the subfamilies *Densovirinae*, which target insects and other arthropods, and *Parvovirinae*, which target vertebrates. The *Parvovirinae* family is further subdivided based on the combination of characteristics such as the ability to autonomously replicate, the packaging of single sense strands of DNA, and the number of promoters in the genome. The resulting genera are the Parvovirses, Erythroviruses, Amdoviruses, Bocaviruses, and Dependoviruses, of which AAV is a member.

Dependoviruses, so named for their dependency on a helper virus for replication, were originally comprised only of the AAVs. Phylogenetic analysis however has now placed the autonomously replicating goose and duck parvoviruses within this genus, where they cluster with avian AAV

(Tattersall, 2006). Additionally, 9 serotypes of AAV (AAV1-9) have been isolated from various hosts and are classified as different species within the genus. While not all members of the dependovirus family share the characteristic of helper dependency, shared characteristics include the packaging of equal numbers of plus and minus strands of DNA 4.7-5.1 kb in length, the use of three promoters at positions 5, 19, and 40, and the use of functional polyadenylation signals at the right hand end of the genome (Tattersall, 2006).

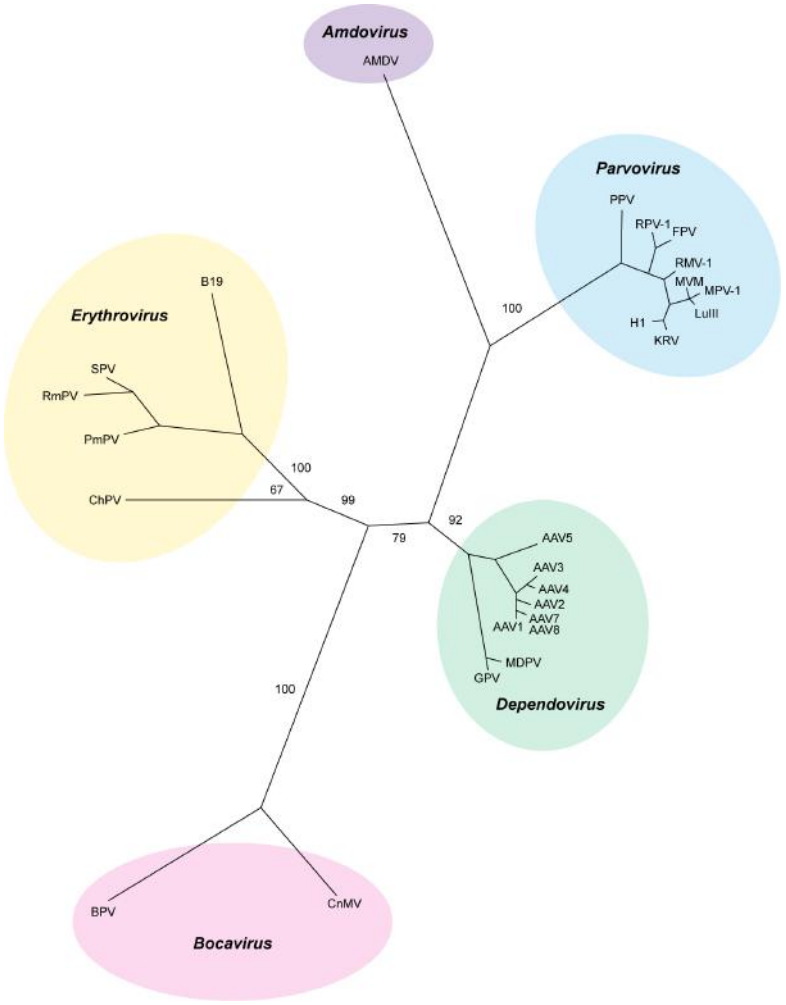


Figure 1. Phylogeny of the subfamily *Parvovirinae*
Phylogenetic relationship between the nonstructural genes of members of the subfamily Parvovirinae. Adapted from Peter Tattersall, 2006. The evolution of parvovirus taxonomy. In: Paroviruses. pg 7

1.1.2.2 The Genome

The AAV genome consists of a 4.7 kb ssDNA molecule flanked by two inverted terminal repeats (ITRs), which are packaged separately as positive or negative sense strands into preformed capsids (Berns and Rose, 1970; King et al., 2001). The negative sense strand alone serves as a template for transcription and is comprised of only two major open reading frames (ORFs) (Rose and Koczot, 1971; Srivastava et al., 1983). The left ORF codes for the nonstructural replication (Rep) proteins, termed Rep40, Rep52, Rep68, and Rep78 based on their apparent molecular weight. These are generated through the use of two different promoters, p5 and p19, and alternative splicing (reviewed in section 1.1.3.6), resulting in four structurally related proteins that together orchestrate virtually every aspect of the viral life cycle. The p40 promoter in the right ORF drives the expression of the three structural capsid proteins – VP1, VP2, and VP3 – which are derived through the use of alternative start codons as well as alternative splicing, and which together form the viral capsid. Additionally, a nested, alternative ORF within the *cap* gene has been described, resulting in the expression of an assembly-activating protein (AAP) necessary for efficient capsid assembly (Sonntag et al., 2010). The AAV genome is flanked by two ITRs comprised of imperfect palindromes that self-anneal to form hairpin-like secondary structures (Lusby et al., 1981). Within the ITRs are two sequences, the Rep binding site (RBS) and terminal resolution site (TRS), which represent the only *cis*-acting signals necessary for viral replication (Samulski et al., 1989). The RBS specifically recruits the large Rep proteins to the origin of replication (McCarty et al., 1994), after which Rep endonuclease activity introduces a nick at the nearby TRS necessary for the resolution of 5' hairpins during replication (Im and Muzyczka, 1990a).

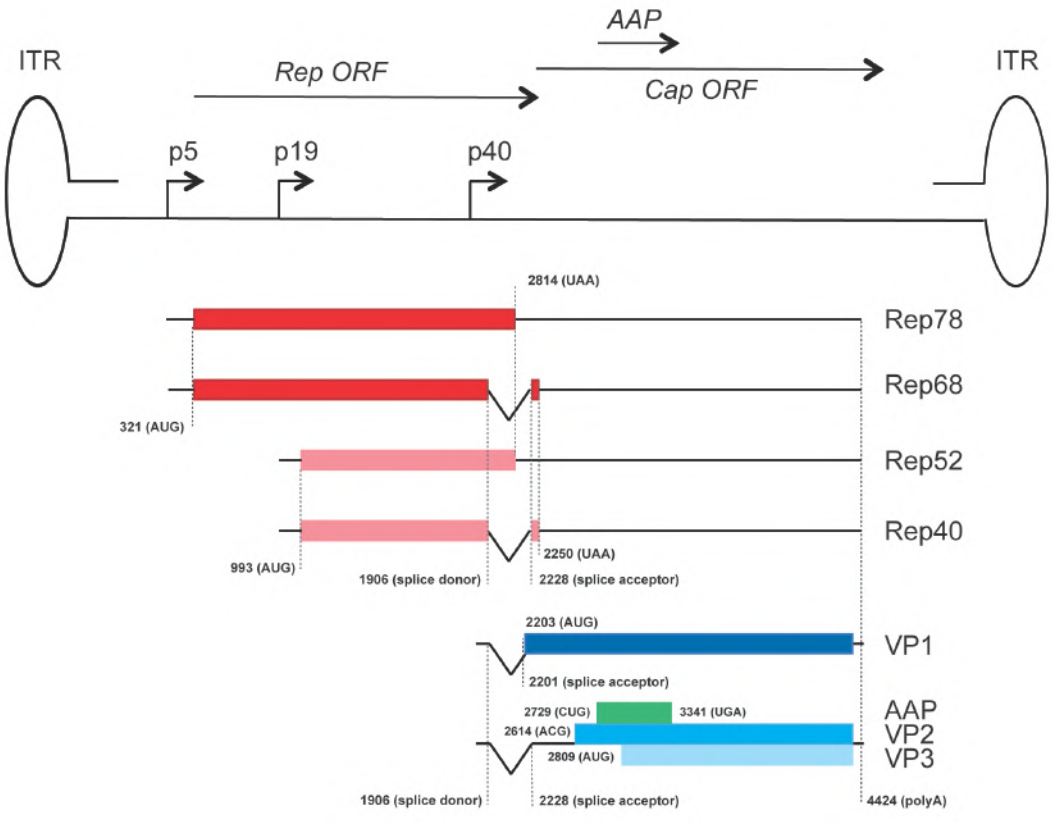


Figure 2. AAV genome, transcriptional units, and gene products
Shown is a schematic representation of the AAV genome, including the various ORFs, as well as the resulting AAV transcripts. Straight right-facing arrows (top of figure) depict the two major AAV ORFs, *rep* and *cap,* which code for the nonstructural replication proteins and the structural VP proteins, respectively, as well as the cryptic ORF, AAP. The ssDNA AAV genome is shown below in black, flanked by two ITRs. Small, bent arrows stemming from the genome represent the three AAV promoters, p5, p19, and p40. Transcripts arising from p5, p19, and p40 are shown in red, pink, and blue, respectively, and the cryptic ORF product is shown in green.

1.1.2.3 The Capsid

The AAV genome is enclosed within a non-enveloped protein coat whose role in the viral life cycle includes not only the protection of the viral genome, but also the mediation of host cell recognition, nuclear trafficking, release of the viral genome upon arrival at the nucleus, and egress from the host cell (Agbandje-McKenna and Kleinschmidt, 2011). Interestingly, recent work with various capsid mutants has also revealed a transcriptional role for the capsid proteins (Aydemir et al., 2016; Salganik et al., 2014). The AAV capsid is ~20-25 nm in size and is composed of 60 protein subunits arranged into a lattice with T = 1 icosahedral symmetry. The three capsid proteins –

VP1, VP2, and VP3 – are 87, 73, and 62 kDa in size, respectively, and are transcribed from the p40 promoter. The use of both alternative splicing and alternative start codons results in three VP proteins that share an identical C-terminal domain, with VP2 and VP3 containing successive amino-terminal truncations. Additionally, translation of VP2 from a weak, unconventional ACG start codon, and VP3 from an internal AUG, supports expression of the capsid proteins in the correct stoichiometry of 1:1:10 (Becerra et al., 1988).

Newly synthesized capsid proteins are targeted to the cell nucleolus by the AAV protein AAP, after which capsid morphogenesis is facilitated by AAP through an unknown mechanism (Sonntag et al., 2010). Preformed capsids are then transported to the nucleoplasm in a Rep-dependent manner for the process of viral genome encapsidation (Wistuba et al., 1997). Several basic regions (BR) within the capsid proteins have been shown to be essential for viral infectivity and propagation. BR1-3, located at the N-terminus of VP1 and VP2, resemble classical nuclear localization signals (NLS) (Grieger et al., 2006; Sonntag et al., 2006). The contribution of each of these BR to nuclear localization remains to be determined, however mutational analysis suggests a stronger contribution of BR3 to the nuclear transport capacity of VP1/VP2 N-termini (Popa-Wagner et al. 2012). BR4, shared by all three capsid proteins, is thought to be essential for virion assembly in the nucleus (Grieger et al., 2006), and a small basic region at the C-terminus of the capsid proteins, BR5, has yet to be analyzed for AAV. More recent work suggests that AAP also contains a redundant, multipartite NLS and nucleolar localization signal (NoLS), and that it plays a critical role in transporting VP3 to the nucleolus for assembly (Earley et al., 2015). In addition, VP1 contains a conserved N-terminal phospholipase A2 (PLA2) domain, which is exposed upon acidification of endosomes after uptake and is essential for endosomal escape of incoming viral particles (Girod et al., 2002; Stahnke et al., 2011).

X-ray crystallography resolution of the capsid structures for AAV serotypes 1 through 9 (DiMattia et al., 2012; Lerch et al., 2010; Miller et al., 2006; Nam et al., 2007; Ng et al., 2010; Padron et al., 2005; Quesada et al., 2007; Walters et al., 2004; Xie et al., 2002) has led to a more profound understanding of the molecular basis for certain variable characteristics of

AAV, such as serotype tropism, capsid stability, and immunogenicity. While all three capsid proteins contribute to the overall structure, VP3 alone appears to constitute the majority of the observable capsid surface. The N-termini of VP1 and VP2 have been described to form globules within the internal capsid at the twofold axis of symmetry, which are externalized upon heat shock or acidification through pores located at the fivefold axes of symmetry (Kronenberg et al., 2005; Sonntag et al., 2006), exposing the putative NLSs and the essential VP1-specific PLA2 domain. The structure of the capsid appears to be mainly formed from a highly conserved region of VP3, consisting of an eight-stranded β-barrel core (Agbandje-McKenna and Kleinschmidt, 2011). These β-barrel cores however are connected by long hypervariable loop structures that intimately interact with those of neighboring VP3 subunits, forming 3-fold peaks that cluster around a depression at the threefold axis of symmetry (Xie et al., 2002), and which account for variability between the serotypes in receptor binding and tissue tropism as well as immunogenicity (Drouin and Agbandje-McKenna, 2013).

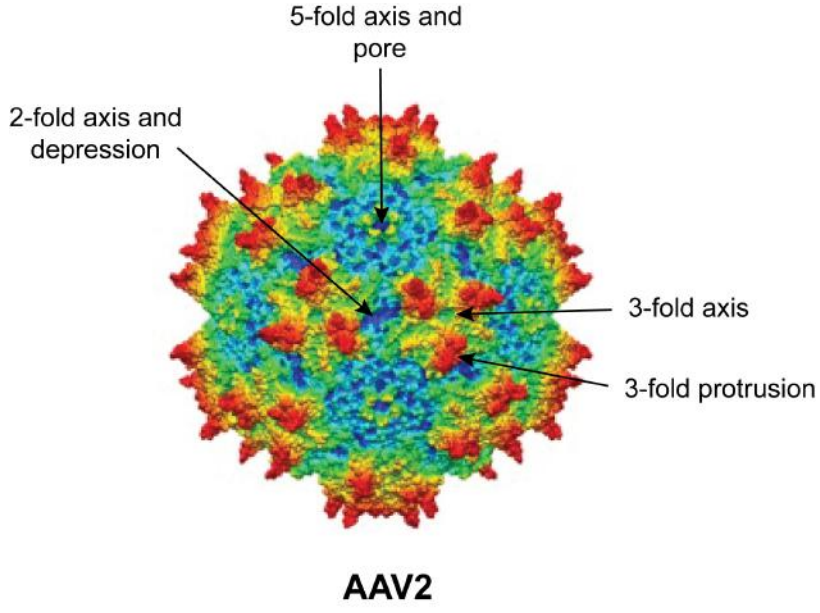


Figure 3. AAV2 capsid structure
Shown is a radially depth-cued surface representation of the AAV2 capsid viewed along the icosahedral 2-fold axis of symmetry. Adapted from (Kailasan et al., 2015).

1.1.3 The Life Cycle

1.1.3.1 An overview

Studies in tissue culture suggest that AAV has evolved a unique, biphasic life cycle in which replication is dependent on both host and helper virus. AAV enters the host cell via endocytosis, after which it is trafficked to the nucleus via the endosomal pathway and enters the nucleus as an intact particle, whereupon the genome is released. In the presence of helper virus coinfection, the AAV genome undergoes several rounds of replication before packaging into preformed capsids and release via cell lysis. In the absence of coinfection, AAV establishes latency either through long-term persistence as an episome, or through preferential integration into specific sites of the host genome. AAV can then be rescued from this latent state through subsequent superinfection by helper virus.

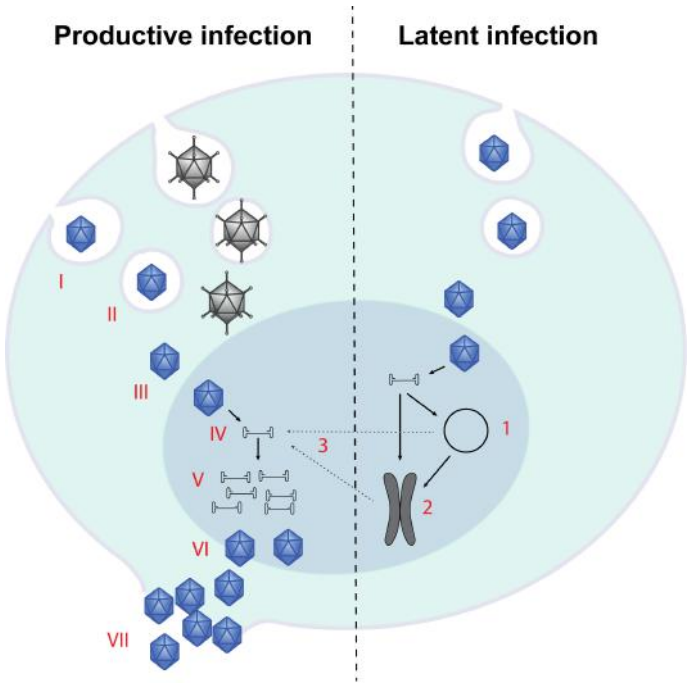


Figure 4. The AAV life cycle
AAV infection can result in either a productive infection in the presence of helper virus (adenovirus pictured here, not to scale) or a latent infection in its absence. The AAV particle enters the cells by receptor-mediated endocytosis **(I)** and is trafficked to the nucleus through the endosomal system **(II)** to the nuclear membrane **(III)**. After tanslocation into the nucleus, viral DNA is released **(IV)** and undergoes several rounds of replication **(V)** before being packaged into preformed capsids **(VI)** and released from the cell **(VII)**. In the absence of helper virus, the AAV particle reaches the nucleus where it can establish latency by forming a stable extrachromosomal episome **(1)** or by integrating into particular sites in the host genome **(2)**. From here, latent AAV can be rescued by superinfection with a helper virus **(3)**.

1.1.3.2 AAV infection in humans

From the very earliest discovery of AAV, it was noted that infection with AAV alone caused no cytopathic effect in various cell types in tissue culture or pathology when inoculated into newborn mice or hamsters (Atchison et al., 1965). It has since become accepted that AAV infection in humans is also non-pathogenic, and no convincing correlation has yet been made between the widespread infection of the human population by AAV and any known disease. AAV DNA has been extracted from various human tissues, including oral, gastrointestinal, and genital tissues (Bantel-Schaal and Hausen, 1984; Blacklow et al., 1967), and seroepidemiological studies have shown that approximately 80% of the human population has detectable levels of antibodies against various AAV serotypes (Blacklow et al., 1968, 1971; Calcedo et al., 2009; Georg-Fries et al., 1984; Mayor et al., 1976; Rosenbaum et al., 1971). By 10 years of age, more than 60% of the population presents with neutralizing antibodies to AAV serotypes 1-3, which generally persist into adulthood. The lack of any detectable pathology associated with AAV infection however presents a real obstacle to the study of the relationship between AAV and its host.

Despite the lack of pathology, certain interesting clinical correlations have been made with respect to a potential protective effect of AAV infection against cervical cancer, although this has not been conclusively established. Epidemiological studies found that 85% of women with normal cervical cytology were seropositive for AAV, as compared to only 14% of women with cervical cancer (Mayor et al., 1976). In addition, it was found that antibody titers were 3-fold higher in healthy women than in women with cervical cancer and that the risk of cervical cancer was inversely correlated with titers of anti-AAV2 IgG in serum (Georg-Fries et al., 1984; Smith et al., 2001). This putative protective effect is posited to result from the helper-dependent life cycle of AAV, in which AAV replication ensures the specific lysis of cells infected with deleterious viruses, such as adeno-, herpes-, and papilloma-virus. This question remains unresolved however, as several other studies have failed to find any correlation between AAV seropositivity and cervical cancer (Odunsi et al., 2000; Strickler et al., 1999).

The lack of a correlation between AAV infection and any pathology has also hindered the study of the viral life cycle *in vivo*. Several groups have attempted to identify target tissues in humans. While AAV has been readily detected in adenovirus isolates from respiratory and gastrointestinal tissues (Blacklow et al., 1967), as well as in the genital tracts of HSV-infected women (Friedman-Einat et al., 1997), these routes of infection likely represent only the productive, helper-assisted phase of the AAV life cycle. Determining the potential reservoir tissue for latent AAV has proven much more difficult. So far, the only tissue found to have high frequencies of AAV is muscle, where 17% of random biopsies tested positive (Tezak et al., 2000). It is also still unclear in what form (i.e. integrated or episomal) AAV establishes latency *in vivo*. Although there is compelling evidence from studies in tissue culture that AAV integrates in a site-specific manner into *AAVS1* in human chromosome 19, one study attempting to characterize genomes isolated from human tissues found that the majority of AAV DNA existed as circular double-stranded episomes (Schnepp et al., 2005).

In agreement with its nonpathogenic nature, AAV is also associated with relatively mild immune toxicity, characterized mostly by a humoral response (Chirmule et al., 1999) initiated via the complement pathway (Zaiss et al., 2008). While the relevance of this response for wtAAV infection is unknown, the high seroprevalence for AAV antibodies in the human population does have implications for AAV gene therapy. A study concerning the prevalence of pre-existing neutralizing antibodies (nAb) to various AAV serotypes showed a high prevalence of nAb to AAV2, followed by AAV1 (Calcedo et al., 2009). This presents an obvious problem for AAV gene therapy in patients who have pre-existing NAB, or alternatively those who may need more than one treatment. Additionally, in one clinical trial, in which liver-directed AAV vectors expressing human Factor IX (AAV-hFIX) were used to treat Hemophilia B, CD8+ T cell responses to AAV capsid in one patient who had received a high dose of vector led to transient and self-limited liver transaminitis and an associated decline in hFIX levels (Manno et al., 2006; Mingozzi et al., 2007). These results highlight the importance of intensifying our efforts to further elucidate the biology and immunology of AAV infection so as to enable the continued development of AAV vectors

with increased bioactivity, which could translate to therapeutic doses low enough as to greatly reduce or even avoid immunological responses.

1.1.3.3 Entry and Trafficking

Much of our current understanding of AAV entry and trafficking is derived from studies using rAAV vectors of various serotypes. As a result, there remains a degree of ambiguity with respect to the particular receptors and trafficking pathways used by wtAAV during the natural infection process. In addition, studies have been performed with tissue culture adapted strains of rAAV, whose capsids have often undergone changes in order to more efficiently target the particular receptors expressed on immortalized tissue culture cell lines. This notion is supported by the finding that tissue culture adapted rAAV2 readily binds heparan sulfate proteoglycans (HSPG), while isolates extracted from human tonsils do not (Chen et al., 2005). It is highly likely that AAV in fact exploits a combination of alternative pathways depending on target tissue, serotype, and the presence or absence of helper virus.

As is the case with many other viruses, AAV uses both a primary attachment receptor, generally consisting of carbohydrate moieties such as glycoproteins or proteoglycans, as well as a secondary proteinaceous receptor for efficient binding and endocytosis. Various attachment receptors have been identified. For AAV2 and AAV3, interactions between two arginine residues at positions 585 and 588 with the proteoglycan HSPG have been shown to mediate cellular attachment (Kern et al., 2003; Opie et al., 2003; Summerford and Samulski, 1998; Summerford et al., 1999). Different linkages of sialic acids provide the equivalent attachment functions for AAV4 and AAV5 (Kaludov et al., 2001; Walters et al., 2001; Wu et al., 2006), while galactose has been shown to serve as an attachment receptor for AAV9 (Shen et al., 2011). There is evidence to suggest that these primary attachment receptors may also induce conformational changes to the capsid, increasing affinity for the secondary entry receptor (Levy et al., 2009). Entry receptors identified so far include fibroblast growth receptor 1 (FGFR-1) and αVβ5 integrin for AAV2, platelet derived growth factor receptor (PDGFR) for AAV5, and epidermal growth factor receptor (EGFR) for AAV6 (Di Pasquale

et al., 2003; Qing et al., 1999; Weller et al., 2010). Recently however, a ground-breaking study in which haploid cell lines were used to screen for essential proteins in AAV2 infection identified an uncharacterized type I transmembrane protein, KIAA0319L, as a universal receptor for AAV (Pillay et al., 2016). The investigators were able to show that this protein was not only essential for infection of a wide range of mammalian cell types by all AAV serotypes tested, but also that *KIAA0319L* $^{-/-}$ mice were resistant to infection by AAV. This protein has since been termed the AAV receptor (AAVR).

Following attachment, AAV is internalized in a rapid and efficient manner through various mechanisms of endocytosis. Roles for clathrin-mediated endocytosis, macropinocytosis, and Clathrin-Independent Carriers/GPI-Enriched Endocytic Compartment (GLIC/GLEEC) endocytosis have been demonstrated in mediating cellular uptake of AAV (Bartlett et al., 2000; Duan et al., 1999; Nonnenmacher and Weber, 2011; Sanlioglu et al., 2000). Additionally, multiple pathways appear to occur simultaneously as inhibition of both clathrin- and GLIC/GLEEC-mediated pathways was shown necessary to fully block viral entry (Nonnenmacher and Weber, 2011). Several lines of evidence suggest that these different pathways lead to alternative fates for the viral particle, either delivering the virion successfully to the nucleus via the infectious pathway, or targeting the virion for degradation or export via the non-infectious pathway. Inhibition of clathrin-mediated endocytosis reduced the uptake of rAAV2 while having no effect on transduction, suggesting this pathway in fact represents a dead-end for infection (Nonnenmacher and Weber, 2011). It has also been observed that certain cell types can efficiently internalize viral particles while being non-permissive to transduction due to impaired post-entry trafficking (Di Pasquale and Chiorini, 2006), and *in vivo* experiments performed with various rAAV serotypes also failed to demonstrate any strict correlation between DNA accumulation and transduction (Miao et al., 2000; Zincarelli et al., 2008, 2010). Perhaps most strikingly, inhibition of transcytosis in polarized cells increased transduction by rAAV4 and rAAV5, but not rAAV2 or rAAV6, suggesting these cells were redirecting vectors to two independent pathways (Di Pasquale and Chiorini, 2006).

In contrast to internalization, intracellular trafficking of AAV is inefficient, with only a small fraction of viral particles reaching the nucleus and the majority accumulating in perinuclear compartments (Bartlett et al., 2000). Retrograde transport intermediates are also still not entirely clear. Studies using various lysosomotropic agents have rigorously demonstrated that AAV is transported through the endosomal compartment (Bartlett et al., 2000; Douar et al., 2001; Sonntag et al., 2006), where acidification triggers essential conformational changes to the capsid necessary for downstream nuclear trafficking and, ultimately, infection. These include externalization of the N-termini of VP1 and VP2, which contain the PLA2 domain necessary for endosomal escape, and BR1-4 necessary for nuclear localization (Johnson et al., 2010; Nam et al., 2011; Sonntag et al., 2006). Studies using fluorescently labeled wtAAV and rAAV have shown that viral particles then accumulate at the trans golgi network (TGN) (Bantel-Schaal et al., 2002; Johnson et al., 2010), however the significance of this for infection is still elusive. The observation that inhibition of endosomal acidification more than 90 minutes post infection no longer inhibits transduction would suggest that infectious particles escape from early endosomes (Bartlett et al., 2000). It is possible that the block to transduction observed with inhibition of the TGN (Nonnenmacher and Weber, 2011) reflects the role of the TGN in mediating endosomal acidification through the provision of hydrolases, rather than the need for viral particles to accumulate within it. It is important to note that trafficking experiments are necessarily performed in the absence of helper virus and using high MOIs of AAV. Only a fraction of AAV is ever truly infectious however (Zeltner et al., 2010), meaning that we cannot be sure that an observed particle is actually reflective of the infectious pathway. It may be that infectious particles escape from early endosomes after acidification, and the majority observed to accumulate at the TGN represent dead particles. Furthermore, the lack of helper virus means that these pathways may represent mechanisms specific to latent infection.

AAV nuclear translocation is also still ill defined. Studies using rAAV2 suggest a role for interactions between the AAV2 capsid proteins and importin-β in entry through the nuclear pore complex (NPC), however antagonism of this pathway only led to partial inhibition in nuclear

translocation (Johnson and Samulski, 2014). Intact AAV particles enter the nucleus and initially become sequestered within nucleolar compartments in stable form (Johnson and Samulski, 2009). Subsequent mobilization of virions to the nucleoplasm leads to uncoating and gene expression or genome degradation, the kinetics of which appear to be cell type- and serotype-dependent (Johnson and Samulski, 2009).

1.1.3.4 Gene transcription and regulation

Adeno-associated viruses exhibit a highly compact, overlapping genetic organization in order to maximize the transcriptional potential of their small genomes. The AAV2 genome has three different promoters – p5, p19, and p40 – all transcripts of which contain a single intron just downstream of p40, which uses a 5' splice donor site at nucleotide 1906 and one of two acceptor sites at either nucleotide 2001 (A1) or 2228 (A2) (Carter et al., 1990; Green and Roeder, 1980; Lusby and Berns, 1982), and a single polyadenylation site. Unspliced transcripts of the p5 and p19 promoters encode Rep78 and Rep52, respectively, while spliced transcripts encode Rep68 and Rep40. Alternative splicing to the acceptors A1 or A2 results in the formation of two different isoforms of Rep68 and Rep40, however no difference in function between the alternatively spliced forms of Rep68 or of Rep40 has ever been reported. Transcripts from the p40 promoter that are spliced using the A1 acceptor encode the VP1 capsid protein, while transcripts using the A2 acceptor encode both VP2 and VP3 (Trempe and Carter, 1988a). VP2 is generated from a non-canonical ACG initiation site, and VP3 from read-through to the next available AUG.

AAV gene expression must be tightly regulated due to the anti-proliferative potential of the Rep proteins. This regulation is achieved through a complex interplay between host cellular, helper virus, and AAV Rep proteins. In addition, Rep binding to the viral RBS can serve to both activate and repress transcription (Pereira et al., 1997), depending on the particular RBS in question and the presence or absence of helper virus proteins. In the absence of helper virus coinfection, three elements within the p5 promoter have been shown to be responsible for repression of all three viral promoters – a major late transcription factor (MLTF) binding site at position -80, a YY1

binding site at position -60, and an RBS at position -20 (Chang et al., 1989; Hörer et al., 1995; Kyostio et al., 1995; Kyöstiö et al., 1994; Pereira et al., 1997). Optimal repression of the p5 promoter is dependent on both Rep binding to the p5 RBS as well as the presence of an intact NTP-binding site within the Rep ATPase domain (Kyostio et al., 1995; Kyöstiö et al., 1994). However, the small Rep proteins, which contain only ATPase and helicase activity, are also able to repress both the p5 and p19 promoters (Hörer et al., 1995; Im and Muzyczka, 1992; Kyostio et al., 1994; Smith and Kotin, 1998).

In the presence of helper virus coinfection, interactions between the helper virus factors Ad E1A or HSV-1 ICP0 with YY1 (Geoffroy et al., 2004a; Lee et al., 1995) lead to the activation of p5 transcription by both YY1 and MLTF, possibly through the presence of a YY1 responsive initiator site at position +1 (Seto et al., 1991). Rep bound to the RBS in the p5 promoter continues to have an inhibitory effect on p5 while simultaneously transactivating both p19 and p40 (Kyostio et al., 1995; Weger et al., 1997) through a common mechanism involving interactions between Rep bound at the p5 promoter and the transcription factor Sp1 bound to the p19 and p40 promoters (Lackner and Muzyczka, 2002; Pereira and Muzyczka, 1997a). These interactions lead to the formation of complexes that bring transcription factors in the p5 promoter, such as YY1, into close proximity with p19 and p40. Additionally, binding of Rep to the RBS in the viral ITR is necessary for the transactivation of all three promoters (Weger et al., 1997).

Pre-mRNA splicing is also a critical determinant for productive AAV infection, which is necessary for regulating the ratio of unspliced to spliced p5- and p19-generated transcripts, as well as the temporal order of appearance of AAV mRNAs throughout the infection cycle. Efficient splicing requires the participation of both the helper virus and the large Rep proteins. Coinfection stimulates splicing of p19- and p40-driven transcripts in a manner dependent on the adenoviral factors E1A, E1B, E2A, E4, and VA RNA, but this is only manifest in the presence of Rep bound to the RBS (Qiu and Pintel, 2002). Rep-mediated enhancement of the relative levels of spliced RNA is inversely correlated with the distance between the promoter and the intron of the affected transcription unit, meaning that p40 transcripts are spliced with the most efficiency, and p5 transcripts the least (Qiu and Pintel,

2002). There is also a temporal order of appearance for the different AAV mRNAs (Labow et al., 1986; Mouw and Pintel, 2000; Trempe and Carter, 1988b). Unspliced p5 transcripts appear first upon activation of p5, but p19-generated transcripts ultimately accumulate to levels greater than those from p5, creating a negative feedback loop on p5 activity. By late infection however, p40 transcripts dominate. Finally, the percentage of each spliced mRNA increases by varying degrees during the course of infection (Mouw and Pintel, 2002), a phenomenon that might be explained by the aforementioned correlation between splicing efficiency and the distance of the RBS from the affected intron.

1.1.3.5 Genome replication

AAV DNA replication occurs via a mechanism often referred to as rolling hairpin replication (Straus et al., 1978; Tattersall and Ward, 1976; Ward, 2006), a variation of the rolling circle replication (RCR) mechanism used by numerous single-stranded phages, eukaryotic viruses, and bacterial plasmids (Chandler et al., 2013). Upon release from the viral capsid, self-annealing of the ITRs provides the heteroduplex structure necessary to prime initiation of unidirectional strand synthesis (Figure 5A, B), resulting in a full-length duplex genome in which one end is in closed hairpin conformation (Figure 5C). Binding of the RBS and subsequent nicking at the TRS by the large Rep proteins (Figure 5D) generates the new free base-paired 3'-OH necessary for replication through the initiating ITR, a process termed terminal resolution (Figure 5E). The result is a linear, double-stranded replication intermediate, in which the two right ITRs can self-anneal (Figure 5F) to initiate another round of replication by strand displacement (Figure 5G). Each cycle of replication yields two possible products: a double-stranded full length AAV genome containing a closed hairpin, and a single-stranded full-length displacement product (Figure 5H). In addition, if a new round of replication from the 3' terminus reaches a closed hairpin before it has been resolved, duplex dimers in head-to-head or tail-to-tail conformation can occur. This mode of replication leads to the inversion of part of the ITR sequence during each round of replication, leading to two alternative conformations designated flip and flop (Lusby 1980).

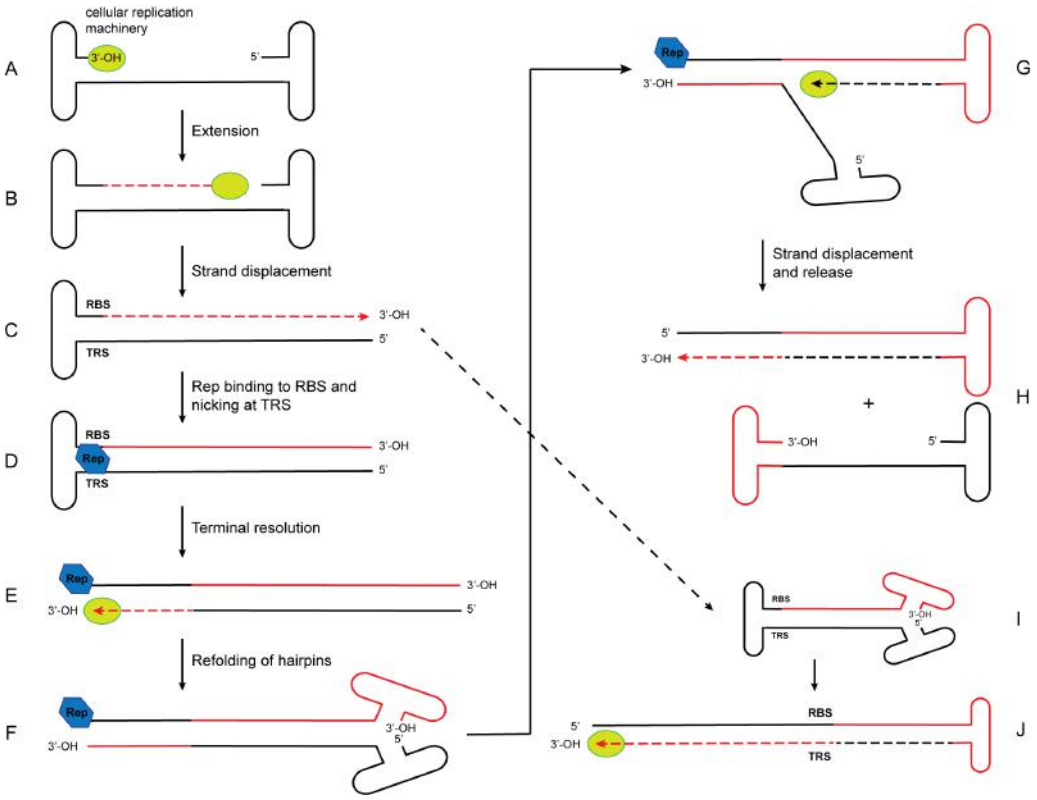


Figure 5. Rolling hairpin replication of AAV
The free 3'-OH provided by the self-annealed ITR serves as a primer **(A)** for DNA replication by the cellular replication machinery **(B)**. The right ITR is replicated by strand displacement, generating a double stranded molecule covalently linked by the left ITR **(C)**. Rep binds to the RBS in the left ITR and generates a nick at the TRS located on the opposite strand **(D)** to allow for terminal resolution of the left ITR **(E)**. The complementary ITRs at the right end of the genome refold **(F)**, generating a new 3'-OH that can be used for another round of replication by strand displacement **(G)** that generates a new single stranded AAV genome and a double stranded molecule covalently linked by the right ITR **(H)**. These molecules could be used for new rounds of replication or alternatively could serve as packaging templates. If refolding of the right ITR occurs prior to terminal resolution of the left ITR **(I)**, a dimeric replication intermediate can be formed **(J)**.

The biochemical interactions between the Rep proteins and the ITRs necessary for replication are extensive. The large Rep proteins contain an origin binding domain (OBD) necessary for substrate recognition, which recognizes and binds the RBS within the ITR (Chiorini et al., 1995; Davis et al., 2000; McCarty et al., 1994; Snyder et al., 1993). Upon binding, the ATP-dependent helicase activity of the Rep proteins unwinds the neighboring region of the ITR containing the TRS (Brister and Muzyczka, 1999; Davis et al., 2000; Im and Muzyczka, 1990a), creating the single-stranded template that is necessary for Rep endonuclease activity. Rep proteins then nick the TRS in a site- and strand-specific manner (Brister and Muzyczka, 1999;

Davis et al., 2000; Im and Muzyczka, 1990a; Smith and Kotin, 2000), becoming covalently linked to the 5' side of the cleavage site in the process (Prasad et al., 1997; Wistuba et al., 1997). Evidence suggests that Rep proteins remain associated with the viral genome for up to 8 hours and may be involved in the encapsidation process, although this is still not entirely clear. Rep helicase activity is then again necessary for the unwinding of the newly freed 3' ITR (Brister and Muzyczka, 1999; Davis et al., 2000; Snyder et al., 1993). The actual hairpin structure of the ITR itself appears to also be important for efficient replication. Hairpins in which regions of the T arms were substituted in such a way as to alter the nucleotide sequence without affecting formation of the intact hairpin structure were still able to support efficient viral replication (Bohenzky and Berns, 1989; Lefebvre et al., 1984).

As AAV can exploit various rather disparate helper viruses, it is thought that, rather than provide specific factors for replication, these helper viruses instead create a cellular environment conducive for replication. It has been shown in *in vitro* reconstitution assays that the minimal protein complexes necessary for AAV DNA replication are replication factor C (RFC), proliferating nuclear antigen (PCNA), the minichromosome maintenance (MCM) complex, and the cellular polymerase δ (Nash et al., 2007, 2008). As with other viruses, AAV replication occurs in foci, or replication centers (RC), which colocalize with an array of cellular and helper virus proteins (Vogel et al., 2013). The particular factors involved are altered depending on the requirements of the different helper viruses, but also depending on the specific function of the RC at a point in time. In the case of herpesvirus, DNA replication and genome packaging is localized to the viral RCs, however these two functions occur in distinct nuclear compartments during adenovirus infection (Vogel et al., 2013). The site of genome packaging for AAV is still not entirely clear, and it may be that AAV RCs temporally and spatially evolve over the course of infection in parallel to the changes in Rep and Cap protein expression.

1.1.3.6 Helper functions

Productive AAV replication is dependent on coinfection of its targeted cell by an unrelated helper virus – such as adeno-, herpes-, and papilloma-

virus (Ad, HSV-1, and HPV) – which provides essential factors necessary to both alter the cellular milieu such that it becomes conducive to viral replication, as well as to directly support AAV transcription and DNA replication. The particular subsets of viral proteins necessary to support AAV replication have been extensively studied and defined for the two major AAV helper viruses, Ad and HSV-1. The minimal set of Ad proteins necessary for AAV replication is comprised of E1A, E1B55K, E2A, E4orf6, and viral associated RNA (VA RNA) (Fisher et al., 1996; Samulski and Shenk, 1988). Interactions between Ad E1A and YY1 bound to the AAV p5 promoter are necessary to relieve repression of p5 and to initiate transcription of the large Rep proteins (Chang et al., 1989; Lewis et al., 1995; Shi et al., 1991; Weitzman et al., 1996). E1B55K and E4orf6 act as a complex to enhance AAV replication through the targeted proteasome-mediated degradation of the Mre11/Rad50/Nbs1 (MRN) complex, and DNA ligase IV (Baker et al., 2007; Stracker et al., 2002). The MRN complex is a primary component of the DNA damage and repair pathway, which has been shown to be recruited to AAV ITRs, whereupon it inhibits second strand synthesis and promotes the concatamerization of replicating genomes (Cervelli et al., 2008; Schwartz et al., 2007). E2A encodes a ssDNA binding protein that colocalizes with AAV RCs and which stimulates the processivity of AAV replication *in vitro,* possibly by helping to maintain in solution the extensive lengths of displaced ssDNA that are inherent to the mechanism of rolling circle replication (Ward et al., 1998). VA RNA is a small, non-coding RNA produced abundantly during Ad infection whose role is to enhance the translation of Ad proteins. Inhibition of the kinase PKR by competitive binding of VA RNA prevents the phosphorylation of the translation initiation factor eIF2α, thereby overcoming the shutdown of protein translation often associated with viral infection, and thus also allowing for the efficient translation of AAV mRNA (Nayak and Pintel, 2007).

In contrast to the helper functions of Ad, which predominantly modulate AAV gene regulation, helper functions provided by HSV-1 relate mostly to DNA replication. The minimal subset of HSV-1 proteins initially defined as being sufficient to support AAV replication were UL5, UL8, and UL52, which together form a helicase/primase complex, and the ssDNA

binding protein ICP8 encoded by the UL29 gene (Stracker et al., 2004; Weindler and Heilbronn, 1991), which performs a similar function in AAV replication as the ssDNA binding protein of Ad. In the context of AAV latency, it was later shown that the proteins ICP0, ICP4, and ICP22 act synergistically to stimulate *rep* expression (Alazard-Dany et al., 2009; Geoffroy et al., 2004b). In addition, the HSV-1 DNA polymerase complex (UL30/UL42) was demonstrated to play a role in the induction of AAV DNA replication (Alazard-Dany et al., 2009), and the HSV-1 polymerase alone (UL30) was shown to replicate the AAV genome efficiently *in vitro* (Ward et al., 2001), effectively replacing cellular replication machinery.

1.1.3.7 Capsid assembly and egress

AAV capsid assembly takes place in nucleolar compartments, and so the initial challenge of the assembly pathway is in the nuclear translocation of newly synthesized capsid proteins. The proposed assembly intermediate competent for nuclear transport is a trimer containing VP3 with VP1, VP2, or VP1 and VP2, in which the basic residues of the VP1/VP2-specific NLS are externalized (Agbandje-McKenna et al., 1998; Lombardo et al., 2000; Xie and Chapman, 1996). Once transported into the nucleus, these partially assembled VP complexes accumulate within nucleoli (Johnson and Samulski, 2009; Wistuba et al., 1997), where the process of capsid formation takes place separated from ongoing DNA replication in Rep-containing RCs (Weitzman et al., 1996; Wistuba et al., 1997). As infection progresses, Cap proteins can be seen mobilizing out of nucleoli and into the nucleoplasm in a Rep-dependent manner (Wistuba et al., 1997) where they then colocalize with Rep-DNA complexes (Hunter and Samulski, 1992; Weitzman et al., 1996).

Based on a detailed pulse-chase study of AAV2 virion assembly, a commonly accepted concept of AAV2 assembly has been established in which replicated, ssDNA strand displacement products are packaged into preformed capsids (Myers and Carter, 1980). Further observations that Rep ATPase activity was essential for efficient packaging (King et al., 2001) and that the Rep proteins physically interact with capsid proteins (Dubielzig et al., 1999; Kube et al., 1997; Prasad and Trempe, 1995; Wistuba et al., 1995)

helped to shed more light on the mechanism involved. It is thought that Rep proteins covalently bound to the 3' ITR of replicated and resolved viral genomes (Prasad et al., 1997) together with a 20 bp signal in the D region of the bound ITR (Wang et al., 1997) acts as a sort of initiation sequence for encapsidation by interacting with Rep helicase complexes covalently bound to the capsid (Hölscher et al., 1995) and thus guiding the replicated viral genome to the capsid pore. The Rep helicase complex then serves as a molecular pump to translocate the genome into the capsid in a 3' to 5' direction (Hölscher et al., 1995; King et al., 2001).

The exact mechanism of egress has not been well characterized. It is generally assumed that non-enveloped viruses exit the nucleus by accumulation and eventual cell lysis. Infectious viral particles can be detected in the culture medium prior to the onset of cytopathic effects in the host cells however (Vandenberghe, 2010), suggesting an active transport mechanism also contributes to particle release. Indeed, it has been demonstrated for the related parvovirus MVM that a phosphopeptide in VP2 consisting of serines at position 2, 6, and 10 may act as a non-conventional nuclear export signal (Carreira et al., 2004; Maroto et al., 2000, 2004), and it is possible that a similar mechanism may be relevant to AAV.

1.1.3.8 Latency and site-specific integration

The ability of AAV to establish latency through integration in the absence of helper virus was recognized soon after its discovery, when Detroit 6 cells that had been infected with 250 IU of AAV were found to still be positive for AAV up to 100 passages later (Berns et al., 1975; Hoggan et al., 1972). Initial southern blot analysis of these cells indicated that viral DNA was integrated as head-to-tail concatemers and that the viral ITRs were present at the integration site (Cheung et al., 1980). Similar results were later observed in numerous other cell types (Handa et al., 1977; Laughlin et al., 1986). In order to further characterize integrated AAV DNA, two viral-cellular junctions obtained from the original Detroit 6 cells were mapped and sequenced, revealing that several copies of AAV were integrated in tail-to-tail conformation (Kotin and Berns, 1989). The cellular sequences flanking the AAV insertion sites were later engineered into probes for southern blot and

used to screen a panel of latently infected human cell lines. The results indicated that a specific cellular sequence had been disrupted in 78% of AAV-infected cells, suggesting that integration had occurred in a site-specific manner (Kotin et al., 1990) at a locus which was later sequenced and termed *AAVS1* (Kotin et al., 1992).

The minimal requirements for AAV site-specific integration were subsequently identified as the large Rep proteins, a viral RBS, and a cellular TRS and RBS located in *AAVS1.* The role of the Rep proteins was demonstrated when transfected plasmids containing a *rep* gene and a selectable marker flanked by the AAV ITRs were shown to integrate site-specifically into *AAVS1* (Shelling and Smith, 1994), and this was definitively demonstrated when Rep provided *in trans* was shown to be sufficient to mediate integration of a plasmid containing only a marker flanked by ITRs (Surovsky et al., 1997). The only *cis*-acting signal necessary to mediate site-specific integration is a viral RBS; transfection of plasmids containing either the RBS alone (Surovsky et al., 1997; Young Jr. and Samulski, 2001) or the p5 promoter (Philpott et al., 2002), which contains an RBS (McCarty et al., 1994), was sufficient to mediate integration of ITR-containing plasmids into *AAVS1*. The presence of critical sequences within the cellular integration site was established when AAV infection of cell lines containing stable Epstein-Barr Virus (EBV)-based episomes into which the *AAVS1* pre-integration site had been cloned revealed that a 33-nt sequence containing both a TRS and RBS were necessary and sufficient to mediate integration (Giraud et al., 1994; Linden et al., 1996a, 1996b).

AAVS1 was initially mapped to the long arm of human chromosome 19, at position 19q13.42 (Kotin et al., 1991; Samulski et al., 1991). Detailed analysis of this region confirmed the existence of both an RBS and a TRS within the first 500 bp of *AAVS1* (Urecelay et al., 1995; Weitzman et al., 1994) similar to those found in the AAV genome (Smith et al., 1999; Ward et al., 2001), which were shown to support Rep-mediated integration (Linden et al., 1996b; Weitzman et al., 1994), 1996). It was later found that these signals existed in the 5' UTR of the *PPP1R12C* gene, only 17-nt upstream from the translation initiation start codon (Tan et al., 2001). This gene codes for the protein phosphatase 1 regulatory subunit 12C, also known as Myosin-

binding subunit 5 (MBS85). The *MBS85* gene product is a regulatory subunit of myosin light-chain phosphatase and thus plays a prominent role in the regulation of actin/myosin assembly and disassembly. Furthermore, *MBS85* is ubiquitously expressed in both human and mouse tissues (Tan et al., 2001), and is closely linked to three other genes also essential for the function of actin, *TNN13, TNNT1,* and *EPS8L1* (Dutheil and Linden, 2006). Indeed, the highly gene dense nature of the AAV integration site *AAVS1* initially made it difficult to reconcile the capacity of AAV for latent persistence through integration into this region with its apparent lack of pathology.

Initial PCR-based studies used to specifically locate viral-cellular junctions showed that integration was taking place within the first exon and intron of the *MBS85* gene, and that most junctions were close to the cellular TRS/RBS (Huser et al., 2002; Kotin and Berns, 1989; Samulski et al., 1991). This observation may not be surprising however as the PCR primers used were all specific for this particular region of *AAVS1.* Analysis of the structure of a number of recombinants produced using EBV-based shuttle vectors containing *AAVS1* pre-integration sites (Giraud et al., 1994) as described above revealed several interesting features that would provide a basis for an initial integration model – (1) the observation that viral-cellular junctions occurred circa 1kb downstream of cellular RBS/TRS sites suggested a role for DNA replication in the integration process, (2) evidence for head-to-tail arrangement of integrated DNA suggested the involvement of a circular replication intermediate of AAV, and (3) rearrangement/inversions of flanking cellular sequences implied a series of template strand switches, a major feature of Rep-mediated AAV replication. These observations led to an initial model for integration through replication in which a double-stranded and circular form of AAV is brought into proximity to *AAVS1* via the simultaneous recognition of cellular and viral RBS by large Rep proteins bound to the viral RBS. The linking of viral to cellular DNA by Rep proteins has previously been demonstrated *in vitro* (Weitzman et al., 1994), and circular AAV could result from end-to-end joining or recombination between the two ITRs. Rep then introduces a nick at the *AAVS1* TRS just upstream of the RBS, and DNA synthesis is initiated leading to the displacement of a single strand of *AAVS1*. After limited extension, the extent of which determines the location of the

viral-cellular junction, the elongating strand undergoes a template switch to the displaced strand, resulting in the observed inversion of flanking sequences. Upon reaching the end of the displaced strand – where Rep is assumed to still be covalently bound at the RBS – the elongating strand makes another template switch onto the circular AAV DNA. After replication of the AAV template, the elongating strand again reaches the Rep-bound RBS and makes a final template strand switch back onto *AAVS1*. The single-strand gap created by the inserted AAV sequence is then repaired by cellular machinery. This model explains both the rearrangements of flanking sequences and the variable location of viral-cellular junctions. It also addresses the issue of head-to-tail integrants, as these could only be the result of circular AAV.

Unbiased approaches based on linker-mediated PCR were later used to identify junctions of chromosomal DNA with both the left and right viral ITRs and to study the molecular organization of AAV integration into *AAVS1* (Henckaerts et al., 2009). In addition to commonly reported hallmarks of viral-cellular junctions – such as microhomologies or insertions of unknown sequences at breakpoints – it was revealed that integration events actually occurred at a much greater distance (i.e. >9 kb) from the cellular TRS than previously believed, indicating that extensive DNA replication takes place. Furthermore, the investigators found that *MBS85* sequences at both junctions were present in 5'->3' orientation. Together these observations led to the introduction of a new integration model (Figure 6), in which AAV integration leads to the partial duplication of *MBS85,* potentially maintaining 2 functional copies of the target gene. This study also took advantage of the presence of a conserved, homologous *AAVS1* site in the mouse genome (Dutheil et al., 2004, 2009) to generate latently infected mouse ES cells, which were used to study both the molecular organization and the functional consequences of AAV integration. The fact that Rep-mediated site-specific integration appears to occur without functionally disrupting the target gene infers that novel technologies based on this unique ability of wtAAV may be safe for the genetic manipulation of stem cells. Further studies with human and mouse ES cells harboring a site-specifically integrated rAAV genome expressing GFP revealed that these cells maintained their stem cell

characteristics and that transgene expression remained robust throughout differentiation.

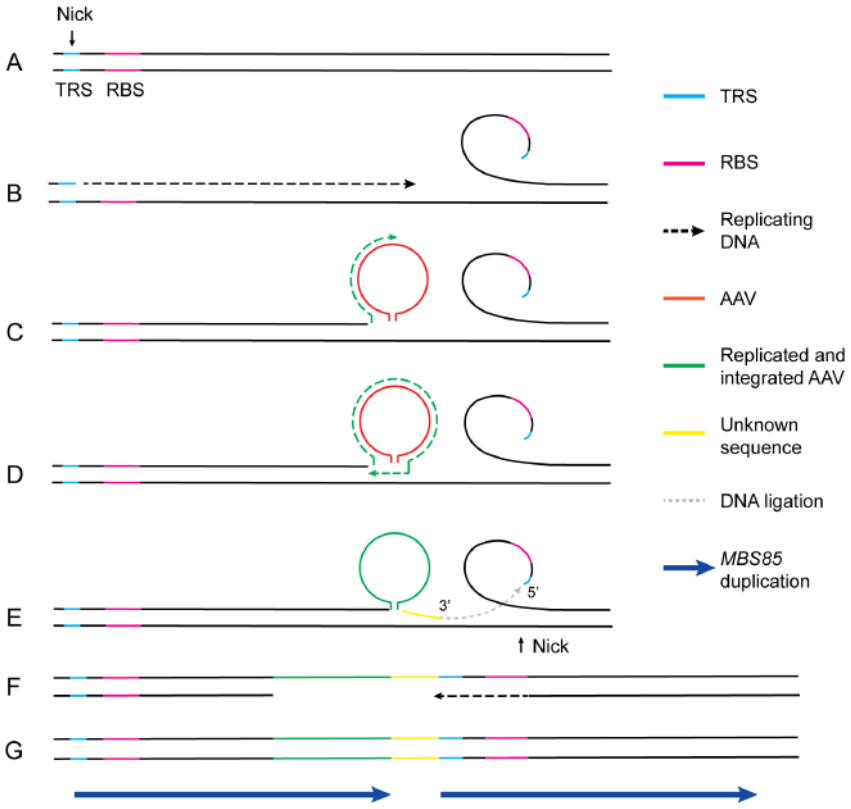


Figure 6. Model for Rep-mediated integration of AAV DNA
(A) Rep binds to the cellular RBS and introduces a strand-specific nick at the cellular TRS in *AAVS1*. **(B)** DNA synthesis is initiated at the nicked TRS, leading to displacement of a single strand of *AAVS1* **(C)** The replication switches template onto AAV. **(D)** Occasional second template strand switch back onto *AAVS1* generates an inverted repeat. **(E)** Ligation between replicated AAV and displaced strand. **(F)** Nicking at the bottom strand stimulates repair of the non-complimentary strand. **(G)** AAV site-specific integration results in partial duplication of *MBS85* sequence. Adapted from (Henckaerts et al., 2009)

Evidence for integration is compelling and has led to the detailed characterization of AAV integration into *AAVS1*. More recently, additional clusters of integration sites, characterized by the presence of RBS motifs, and in some cases, cryptic TRS motifs, have been found throughout the human genome (Hüser et al., 2014; Janovitz et al., 2013; Petri et al., 2015). In addition, regions of open chromatin structure were shown to influence patterns of AAV integration (Hüser et al., 2014). It is important to note however that the majority of observations supporting AAV integration have been obtained in tissue culture experiments, and there is very little data

available with respect to the molecular characterization of latent AAV *in vivo*. AAV DNA sequences have been detected in a variety of tissues, however these studies did not present any data on integration into *AAVS1,* or other integration sites. One study in which sensitive, unbiased PCR techniques were used to characterize AAV DNA obtained from 7 AAV-positive tonsil-adenoid samples showed that the viral genome existed predominantly as a circular double-stranded episome in head-to-tail configuration (Schnepp et al., 2005). Only one sample was shown to contain AAV integrated into chromosome 1. In addition, while it is known for numerous other DNA viruses that latency is regulated through the epigenetic modification of viral chromatin structure, the potential contribution of such mechanisms on AAV latency have yet to be extensively studied. Early evidence that AAV assumes a chromatinized configuration shortly after infection would support a role for epigenetic modification in the establishment and/or maintenance of AAV latency (Marcus-Sekura and Carter, 1983).

1.2 THE AAV2 REP PROTEINS

1.2.1 The Rep proteins and their roles in the AAV life cycle

The particular genomic organization of the *rep* ORF generates four structurally related replication proteins that together orchestrate virtually every aspect of the AAV life cycle. As described in section 1.1.2.2, the use of two promoters and alternative splicing yields four structurally related proteins, termed Rep78, -68, -52, and -40 after their apparent molecular weight, which are comprised of various combinations of three modular domains. The large Rep proteins, Rep68 and Rep78, contain an OBD necessary for DNA binding and site-specific endonuclease activity, while all four Rep proteins share a central ATPase domain containing helicase activity. Additionally, Rep78 and Rep52 share a C-terminal zinc finger (ZNF) domain implicated in several proteins interactions (Figure 7).

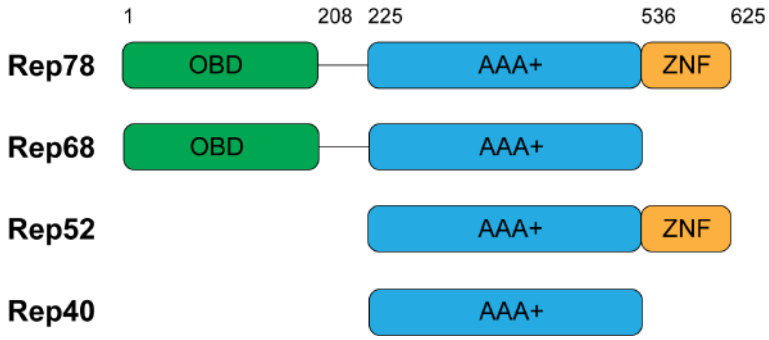


Figure 7. The Rep proteins
Schematic diagram depicting the domain architecture of the four Rep proteins. The OBD, containing DNA binding and endonuclease activities is shown in green; the central helicase domain shared by all four Rep proteins is shown in blue; the C-terminal ZNF domain shared by Rep78 and Rep52 is shown in yellow.

A consequence of this modular structure is that the Rep proteins, in particular Rep78 and Rep68, are versatile and multi-functional proteins with diverse biological functions. The large Rep proteins are essential for virtually every aspect of the AAV life cycle. Binding of the viral RBS by the Rep OBD and subsequent nicking at the TRS forms the basis of terminal resolution during DNA replication, a process necessary to replicate through initiating ITRs and to regenerate 3'-OH groups for new rounds of replication (Ward, 1996). This activity also forms the basis of AAV integration into *AAVS1* as well as subsequent rescue of integrated AAV upon helper virus superinfection (Dutheil and Linden, 2006). All four of the Rep proteins also regulate transcription from the three AAV promoters, as well as helper virus promoters, ensuring minute levels of protein expression during latency, and optimal regulation of lytic protein expression during coinfection. Regulation of certain cellular promoters by the Rep proteins has also been observed and is thought to play a role in maintaining a cellular environment conducive to viral replication. In addition to their role in transcriptional regulation, the small Rep proteins, Rep52 and Rep40, are also essential for the accumulation of single-stranded genomes and their efficient packaging into preformed capsids (King, 2001). Finally, the ZNF domain shared by Rep52 and Rep78 has been implicated in several protein interactions with the potential for disrupting the cell cycle, and so it is possible that these Rep proteins have additional roles in modifying the cellular milieu for viral replication. Numerous other protein

interactions not dependent on the ZNF domain have also been identified; many of these remain poorly defined however, and it is likely that their characterization will illuminate numerous new and significant roles for the Rep proteins in AAV propagation.

1.2.2 Rep domains and enzymatic functions

The OBD is comprised of the first ~200 amino acids shared by the large Rep proteins, Rep78 and Rep68, and contains sequence-specific DNA binding activity as well as strand- and sequence-specific endonuclease activity, conferring upon these proteins the ability to bind and nick DNA substrates containing RBS and TRS, respectively (Chiorini et al., 1994; Davis et al., 2000; Im and Muzyczka, 1990a; McCarty et al., 1994; Owens et al., 1993; Yoon et al., 2001). It shares certain essential features with proteins known to mediate rolling circle replication (RCR), such as the gene A proteins from the bacteriophage ϕX174 and the NS1 proteins of plant geminiviruses (Koonin and Ilyina, 1992, 1993). These include the conserved HUH motif, involved in metal ion binding, and either one or two tyrosine residues shown to be catalytically essential for endonuclease activity (Figure 8) (van Mansfeld et al., 1986; Odegrip and Haggård-Ljungquist, 2001). In addition, the Rep OBD shares structural homology with the DNA-binding proteins SV40 T antigen and papillomavirus E1 (Enemark et al., 2000; Hickman et al., 2002).

The central helicase domain (aa 225-536) shared by all four Rep proteins consists of a motor domain with motifs required for ATPase and helicase activity, as well as a nuclear import signal (Figure 9) (Im and Muzyczka, 1990b; Smith and Kotin, 1998; Wonderling et al., 1995). The AAV Rep helicase belongs to the superfamily 3 (SF3) of helicases, a family found mainly in small DNA and RNA viruses and which includes the helicases of papillomaviruses, poliovirus, and SV40. Like other SF3 helicases, the AAV Rep proteins couple the energy derived from NTP hydrolysis to DNA unwinding in a 3' to 5' direction. The signature of this family is a short stretch of ~100 aa containing the Walker A and B motifs, which comprise the NTP-binding site, the B' box necessary for metal cation binding, and the sensor 1 site. Structural studies of Rep40, which is equivalent to the minimal helicase

domain (Figure 9), and SV40 LTag revealed that these helicases are structurally similar to the AAA+ class of cellular proteins (James et al., 2003) that couple ATP hydrolysis and DNA unwinding.

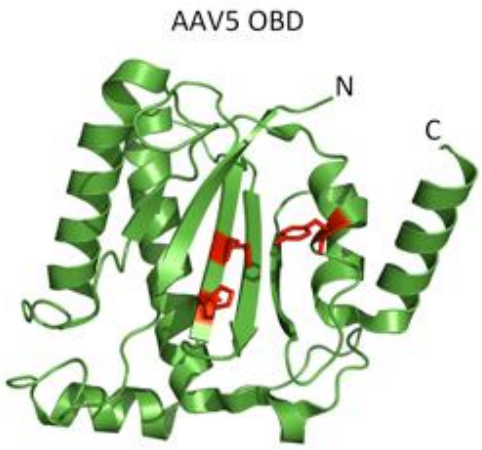


Figure 8. Crystal structure of the AAV5 Rep OBD
The structure shows the five central anti-parallel strands, with the characteristic HUH motif located in the central strand and highlighted in red. This motif is facing the active site tyrosine located in one of the α-helices, also highlighted in red.

The combined activity of the origin-binding and ATPase domains in the large Rep proteins is essential for the processes of genome replication and integration. In both cases, Rep binding of the RBS is followed by nicking at the TRS. As the endonuclease activity of the Rep proteins is specific for ssDNA, this process necessarily occurs in two steps. First, ATP-dependent Rep helicase activity unwinds the TRS in a 3' to 5' direction, resulting in the extrusion of a stem-loop structure and exposing the TRS (Brister and Muzyczka, 1999; Davis et al., 2000). The second step is a strand- and sequence-specific nicking reaction at the sequence 3'-GCT/TGA-5', which is catalyzed by the active site Tyr-156. Transesterifiation of the T/T phosphodiester bonds creates a free 3' OH for subsequent rounds of replication and leaves Rep covalently bound to the 5' thymidine residue in the AAV genome (Brister and Muzyczka, 1999; Smith and Kotin, 2000).

Figure 9. Crystal structure of Rep40 (AAA+ domain)
The structure of Rep40, defined as the minimal helicase domain of AAV Rep, is bi-modular and contains an N-terminal small helical bundle (bottom) and an AAA+-like core characteristic of SF3 helicases.

Both the OBD and ATPase domain are also involved in transcriptional regulation by the Rep proteins via two independent mechanisms - (1) OBD-dependent repression by the large Rep proteins, most likely through steric hindrance of the p5 promoter as a result of binding to the RBS (Dutheil et al., 2014), and (2) ATPase-dependent repression of all three AAV promoters, possibly through interactions between the ATPase/helicase domain of the Rep proteins and components of the cellular transcriptional machinery. Rep68 has also been shown to act as an RNA/DNA helicase (Wonderling et al., 1995), which may provide a mechanism for the observed ATPase-dependent gene regulatory activity of the Rep proteins. Like other AAA+ proteins, the helicase domains of the large Rep proteins interact with one another to form concentration- and ATP-dependent multimers both *in vitro* and *in vivo*, preferentially forming hexamers in the presence of AAV-derived substrates (Mansilla-Soto et al., 2009; Smith et al., 1997). It is thought that these structures are (1) necessary for efficient helicase activity and (2) reflective of the multi-functionality of the Rep proteins. Additionally, the helicase activity of the small Rep proteins, Rep52/40, is essential for the

translocation of full-length, capsid-associated ssDNA AAV genomes into preformed capsids during the process of encapsidation (King et al., 2001).

The C-terminal end of Rep78 (aa 534-607) and Rep52 consists of a short zinc finger motif composed of three CXXC and three CXXH sequences shown to bind zinc *in vitro*. Mutation of either of the repeated motifs abolished zinc binding, suggesting that both of these repeats act together to form a functional ZNF (Hörer et al., 1995). Zinc binding by consecutive CXXC/CXXH motifs leads to the formation of a series of stable, finger-like protrusions that make tandem contacts with their target molecule. Several mammalian proteins involved in chromatin and DNA modification contain CXXC zinc finger domains (Frauer et al., 2011), however no interactions between the Rep ZNF and DNA have been documented. Rather, the Rep ZNF has been implicated in various protein interactions (reviewed in section 1.3.1) with implications for cell cycle regulation. For example, binding and inhibition of the phosphatase Cdc25 by the Rep ZNF domain prevents progression of the cell cycle through the G_1 to S checkpoint. Furthermore, interactions between the Rep52/Rep78 ZNF and the catalytic subunit of protein kinase A (PKA_c) and its close homolog PKrX are thought to interfere with the transcriptional regulation of several proteins involved in cell cycle regulation by interfering with CREB signaling.

1.3 Interactions with the host and helper virus

1.3.1 Rep interactions with cellular proteins

Given the limited repertoire of AAV proteins, the Rep proteins must interact with an array of host cellular factors in order to successfully navigate the complex viral life cycle. Indeed, a multitude of factors involved in DNA replication, transcription, RNA splicing, and the cellular DNA damage response (DDR) have been shown to colocalize with AAV2 replication centers, as well as to copurify with Rep (Vogel et al., 2013). For example, while both the large Rep proteins have the intrinsic DNA-binding and endonuclease activity necessary for terminal resolution and genome integration, in vitro assays have shown that binding of both Rep68 and Rep78 to the cellular ssDNA binding protein replication protein A (RPA) and

nucleophosmin enhances Rep binding to the viral ITR and stimulates endonuclease activity (Bevington et al., 2007; Stracker et al., 2004). Similarly, binding of Rep78 to the non-histone chromosomal protein high mobility group protein 1 (HMG1) stimulates Rep endonuclease activity and ATP hydrolysis and promotes the formation of Rep-DNA complexes (Costello et al., 1997) through the ability of HMG1 to increase DNA flexibility and promote the formation of nucleoprotein complexes.

In addition to their well-characterized role in AAV DNA replication and integration, the Rep proteins both activate and repress transcription from the three AAV promoters (Hörer et al., 1995; Kyostio et al., 1995; Pereira and Muzyczka, 1997a, 1997b; Pereira et al., 1997). Rep-mediated repression has also been observed for various heterologous promoters (Hörer et al., 1995; Labow et al., 1986; Needham et al., 2006). Repression by the Rep proteins is based on two distinct mechanisms, one relying on direct binding of the large Rep proteins to viral (or cellular) RBS, and the other requiring only an intact NTP-binding site within the helicase domain shared by all four Rep proteins, indicating the possible reliance of this pathway on interactions with cellular factors (Dutheil et al., 2014; Kyostio et al., 1995; Kyöstiö et al., 1994). Supporting this notion, Rep is known to interact with various transcription factors (TF) in a repressive capacity. Interactions between the N-terminal domain of Rep78/Rep68 with the transcription factor TATA box binding protein, or TBP, were shown to inhibit the adenovirus major late transcription factor (Hernonat et al., 1998; Needham et al., 2006). TBP is a core factor required for the assembly of the transcription initiation complexes associated with RNA polymerases I, II, and III, and may thus represent one of the most central proteins in transcription initiation. These studies indicate that Rep interaction with TBP may alter the preinitiation complex of RNA pol II, resulting in transcriptional repression (Needham et al., 2006). Similarly, Rep68/Rep78 were shown to bind to the ssDNA binding domain of the transcriptional activator PC4 in an ATP-dependent manner, an interaction suggested to be necessary for the inhibition of cellular transformation by the adenovirus E1A protein, presumably through repression of the associated oncogenes (Weger et al., 1999). Finally, binding of Rep78 to c-Jun downregulated transcription from AP-1 dependent promoters (Prasad et al.,

2003). Certain Rep/TF interactions can also lead to transcriptional activation rather than repression. As described earlier, interactions between the large Rep proteins bound to the AAV p5 promoter and the transcription factor Sp1 bound to the p19 and p40 promoters is thought to be necessary for transactivation of p19 and p40 (Pereira and Muzyczka, 1997a, 1997b) through the formation of a DNA loop that brings essential transcriptional machinery into proximity with p19 and p40.

A number of studies have highlighted the oncosuppressive and antiproliferative properties of the AAV Rep proteins (Hermanns et al., 1997; Winocour et al., 1988). The oncosuppressive effect is thought to be due to the inhibition of tumor-causing viruses by AAV, such as adenoviruses and HPV (reviewed in section 1.3.3). However, the large Rep proteins have also been observed to interfere independently with cell cycle progression by inducing arrest in all cell cycle phases (Saudan et al., 2000) and to induce apoptosis (Schmidt et al., 2000). This is thought to occur via several converging mechanisms. Both Rep68 and Rep78 arrest cells in G_1 and G_2 phases by producing nicks in cellular chromatin and inducing an ATM-mediated DNA damage response (Berthet et al., 2005; Saudan et al., 2000). In addition, it was observed that Rep78 leads to a complete arrest in S phase, a response rarely seen after DNA damage. This activity was later mapped to the C-terminal ZNF domain present in Rep78, which was shown to bind and inhibit the phosphatase Cdc25 (Berthet et al., 2005). The cell cycle is driven by a family of proteins called the cyclin-dependent kinases (CDKs), which are activated in a sequential manner through the removal of inhibitory phosphates by the Cdc25 family of protein phosphatases. This in turn inactivates the tumor suppressor retinoblastoma protein (pRb) through phosphorylation, releasing the transcription factor E2F1 and thus activating expression from E2F1-controlled genes necessary for progression through the G_1 to S phase restriction point. Rep78 binds Cdc25 as a pseudosubstrate, preventing access to its substrates Cdk1 and Cdk2 and therefore preventing phosphorylation/inactivation of pRb (Berthet et al., 2005). This Rep-mediated inhibition of Cdc25 combined with DNA damage induced by the intrinsic endonuclease activity of Rep78 leads to a potent arrest of DNA replication within S phase, characterized by the accumulation

of active, hypophosphorylated pRb (Berthet et al., 2005; Saudan et al., 2000). The importance of this interaction for AAV is underscored by the existence of two supplementary pathways towards the same goal. Rep78 additionally binds E2F1, stabilizing E2F1-pRb complexes and preventing E2F1 release, and interferes directly with expression of E2F1 by binding to and repressing the *E2F1* promoter (Batchu et al., 2001).

A number of viruses are known to regulate their gene expression through the use of *cAMP response elements* (CRE) (Flamand and Menezes, 1996; Kwok et al., 1996; Leib et al., 1991; Nokta and Pollard, 1992). Cellular CRE sites are also involved in the regulation of a variety of cellular processes, including differentiation, proliferation, and apoptosis. CRE sites are regulated by the transcription factor cAMP response element-binding (CREB) protein, which becomes activated when phosphorylated by the cAMP-mediated activation of the protein kinase A (PKA). Rep78 and Rep52 were shown by yeast-two-hybrid experiments to interact via their shared C-terminal ZNF domain with the catalytic subunit of PKA (PKA_c), as well as its close homolog PKrX, and to inhibit PKA- and PKrX-mediated CREB signaling through a protein kinase inhibitor (PKI) like domain in the ZNF domain (Chiorini et al., 1998; Di Pasquale and Stacey, 1998). Inhibition of PKA by PKI plays a central role in the induction of mitosis, and several proteins involved in cell cycle regulation are regulated by CRE in their promoters (Guo et al., 1997; Melendez et al., 1995; Wen et al., 1995; Yoshizumi et al., 1997). Both AAV and MVM contain CRE elements in their p5 and p4 promoters, respectively, and while the function of this site for AAV is unclear, cAMP-dependent stimulation of the CRE of MVM was shown to have a negative regulatory effect (Flotte et al., 1992; Perros et al., 1995). In addition, stimulation of the cAMP pathway, which leads to CREB phosphorylation, strongly enhances both the transcriptional activation and apoptotic properties of the tumor suppressor p53 (Giebler et al., 2000), the cellular accumulation of which induces cell cycle arrest or apoptosis in response to a variety of cellular stress signals, including DNA damage. Interference of CREB signaling by the Rep proteins may thus provide another way to interfere with the cellular DDR.

Finally, while the contribution of chromatin structure to AAV transcription is not yet clear, it is known for various other DNA viruses that remodeling of incoming viral genomes must occur in order to create an active template for transcription and replication (Knipe et al., 2013). For Ad and HSV-1, this is a function provided by the template activating factor I/Set oncoprotein (TAF1/Set) complex, a histone chaperone that acts in complex with the acidic nuclear proteins 32A and 32B (ANP32A and ANP32B). Rep68 was shown to interact with ANP32B via its N-terminal domain, subsequently recruiting the remaining components of the complex, ANP32B and TAF1/Set (Pegoraro et al., 2006). Manipulation of these various components demonstrated a requirement for TAF1/Set activity in the replication of both WT and recombinant AAV (Pegoraro et al., 2006), suggesting that a mechanism similar to what has been demonstrated for Ad and HSV-1 may be necessary to prepare the AAV genome for transcription and replication.

1.3.2 Cellular DNA damage response to AAV

DDR signaling and repair pathways are controlled by the family of PI3K-related protein kinases, comprised of ATM, ATR, and DNA-PK (Linn, 2004). DNA double strand breaks (DSB) in particular represent one of the most dangerous forms of DNA damage, and several independent pathways are in place to repair DSBs. During cellular S/G_2 phase, when sister chromatids are easily available, recruitment of ATM to DSBs by the DNA sensor complex MRN leads to ATM-mediated phosphorylation of the downstream effectors Chk2 and p53, cell cycle arrest, and repair via homologous recombination (HR). ATR is recruited to ssDNA coated in ssDNA-binding protein RPA, a process which naturally occurs either upon ATM-mediated resection of DSBs or at stalled replication forks, whereupon the activation of the downstream effector Chk1 also leads to HR pathways. In the absence of available homologous DNA however, DNA-PK is recruited to DSB by its regulatory proteins Ku70/Ku80, where it mediates repair via non-homologous end joining (NHEJ) through the recruitment of DNA ligase IV.

Both the structure of incoming viral DNA as well as viral replication intermediates have been shown to mediate the recruitment of various components of the cellular DDR pathway. In the absence of a helper virus,

single-stranded AAV DNA has been shown to trigger damage signaling via ATR activation resembling the cellular response to a stalled replication fork (Jurvansuu et al., 2005). It is also worth noting that expression of the large Rep proteins alone can induce a mild activation of ATM and H2AX (Glauser et al., 2010; Schwartz et al., 2009). This is believed to result from the nicking, DNA-binding, and helicase activities of Rep (Berthet et al., 2005; Glauser et al., 2010; Schmidt et al., 2000), however the possibility cannot be excluded that interactions between Rep and cellular proteins are involved. Indeed, Rep proteins have been shown to interact with various cellular proteins involved in the DDR, including RAD50, RPA, γ-H2AX, NBS1, DNA-PK_{cs}, and p53 (Collaco et al., 2009; Nash et al., 2009; Schwartz et al., 2009). In the context of AAV and Ad5 coinfection, AAV replication elicits a robust DDR as demonstrated by the phosphorylation of ATM, Chk1, Chk2, RPA32, and H2AX (Collaco et al., 2009; Schwartz et al., 2009). Studies using rAAV plasmids in the presence of Ad5 and Rep *in trans* suggested that AAV replication is required to induce the DDR (Schwartz et al., 2009). Using chemical inhibitors as well as ATM- and DNA-PK-deficient cell lines, the investigators were able to demonstrate that DNA-PK is the primary mediator of damage signaling in response to AAV replication (Collaco et al., 2009; Schwartz et al., 2009). Supporting this, immunofluorescence analysis confirmed the recruitment of DNA-PKcs, Ku70, and Ku80 to AAV RCs in both Ad and HSV-1 supported AAV RCs (Schwartz et al., 2009; Vogel et al., 2012). In AAV and HSV-1 coinfected cells however, signaling is mediated by both ATM and DNA-PK (Vogel et al., 2012). Furthermore, the changing prevalence of the different AAV replication intermediates within AAV RCs is likely to affect the composition of signaling proteins involved.

The consequences of DDR signaling for AAV replication are still under debate. Genotoxic agents of various forms have been shown to enhance AAV replication, even leading to helper-independent *rep* expression and minimal levels of DNA replication (Nicolas et al., 2012; Yakobson et al., 1987, 1989; Yalkinoglu et al., 1988), however the exact mechanism for this has never been understood. It has been suggested that AAV may need repair factors for processing of replicating genomes; it has also been suggested that inducing a DDR effectively sequesters these proteins away

from AAV RCs, allowing for more efficient replication. In the case of AAV and Ad coinfected cells however, depletion of DNA-PK leads to enhanced AAV replication (Collaco et al., 2009), presumably because DNA-PK would result in the NHEJ-mediated concatamerization of replicating genomes, thus simultaneously removing them from the pool of replicating DNA and preventing encapsidation. Depletion of ATM or ATR however has the opposite effect (Collaco et al., 2009). The reason for this is not clear, and it may be that phosphorylation of a downstream target of these kinases plays a role in AAV replication. Conversely, studies using self-complementary rAAV vectors have shown that DNA-PK is necessary to form the stable episomes that lead to long term transgene expression (Cataldi and McCarty, 2010; Choi et al., 2006). It is possible that a similar mechanism is true for the establishment of stable, circular episomes of latent wtAAV. Furthermore, ATM/ATR activation has a deleterious effect on episomes derived from single-stranded rAAV vectors (Cataldi and McCarty, 2010). As DNA resection is inherent to the ATM/ATR pathway, it is possible that processing of ssDNA genomes results in ITRs that can no longer undergo productive concatamerization.

During AAV and Ad coinfection, degradation of the MRN complex, which is the principal mediator of ATM activation, is an essential helper function for AAV replication. Studies using complementation of these proteins as well as Ad E1B55K/E4orf6 mutants showed that degradation of Mre11 enhances wtAAV replication and rAAV transduction (Cataldi and McCarty, 2010; Cervelli et al., 2008; Choi et al., 2006; Schwartz et al., 2007). Components of the MRN complex are recruited to the viral ITRs in a manner dependent on the hairpin structure, where they lead to silencing of the genome (Cataldi and McCarty, 2013; Lentz et al., 2015). This effect seems to be independent from the activation of downstream effectors, such as ATM, and may instead rely on some consequence of the physical interaction (Lentz et al., 2015). Interestingly, in HSV-1 supported AAV replication, all components of the MRN complex are recruited to AAV RCs without any apparent inhibitory effect (Nicolas et al., 2010; Vogel et al., 2012). It has been suggested that the utilization of the HSV-1 polymerase for AAV

replication may somehow bypass the observed inhibitory effects of the MRN complex (Vogel et al., 2013).

1.3.3 AAV interactions with helper virus

In addition to its effects on the proliferation of the host cell in the absence of helper virus, AAV also exerts a profound effect on the replication of its helper viruses and on cellular transformation. AAV is dependent on various helper factors for its replication (reviewed in section 1.1.3.5), however competition for space and resources within the host cell means that AAV must limit the replication of its helper virus without affecting the synthesis of the helper factors required for AAV replication. It has long been observed that AAV interferes with Ad propagation during coinfection, leading to up to 100-fold reduction in the production of infectious Ad particles, and up to a 10-fold decrease in Ad DNA replication (Carter et al., 1979; Laughlin et al., 1979a; Weitzman et al., 1996). Evidence that AAV may regulate Ad propagation both at the transcriptional level and by interfering with Ad-mediated cellular transformation may begin to explain the balance between AAV's requirement for adenoviral early proteins, and the ongoing competition for cellular space and resources, as well as helper virus factors.

Transcriptional regulation of Ad gene expression by AAV has been demonstrated on various levels. As described in section 1.3.2, AAV interferes with CREB-signaling through the inhibition of the cAMP-responsive kinase PKA. CRE sequences regulate the expression of numerous viral genes (Gilchrist et al., 1996), and the cAMP response pathway components are often used to control switches from early to late or latent to lytic replication modes (Leib et al., 1991; Leza and Hearing, 1989; Pei and Berk, 1989; Xia et al., 1996). Ad has been shown to be dependent on protein kinase activation both for transport into the nucleus, and for E1A-mediated transactivation of *E1A*, *E3*, and *E4*, all of which contain PKA-responsive CRE within their promoters (Leza and Hearing, 1989; Suomalainen et al., 2001). While mutated AAV genomes lacking the PKI-like domain necessary for PKA inhibition could replicate efficiently in the presence of Ad helper plasmids, these mutant variants were outcompeted in the presence of WT Ad, suggesting that this pathway is necessary to preserve AAV replication fitness

during Ad coinfection by selectively targeting Ad gene expression (Pasquale et al., 1998). This effect was particularly profound at low MOI, indicating PKA inhibition may be an absolute necessity in the context of limiting concentrations of AAV.

Several studies demonstrated a role for the Rep proteins in the direct transcriptional regulation of Ad gene expression. Cotransfection of HeLa cells with Ad-luciferase and AAV *rep* expressing plasmids showed that Rep78 could repress expression from all five Ad early gene promoters when in the presence of over-expressed E1A, indicating that Rep could interfere with E1A-mediated transactivation of E1B, E2, E3, and E4 (Jing et al., 2001). This effect could conceivably result from the aforementioned inhibition of CREB signaling. All of the Ad early gene promoters also contain TATA or noncanonical TATA boxes however. As mentioned in section 1.3.1, Rep has been shown to physically interact with the TATA binding protein, TBP, which suggests that the promiscuous repression of Ad early gene expression by Rep may be a result of Rep-specific interactions with TBP, preventing its activation by E1A (Jing et al., 2001). In addition, both Rep68 and Rep78 were shown to inhibit *E2A* expression by directly binding to a sequence between the *E2A* TATA region and mRNA start site that is homologous to the RBS' in the AAV ITR (Casper et al., 2005; Jing et al., 2001; Nada and Trempe, 2002). The presence of two binding sites for the transcription factor E2F also suggests that interactions between Rep, TBP, and E2F1 (also shown to interact with Rep) bound to *E2A* may serve to stabilize the weak Rep-DNA interaction.

Dose-response and temporal analyses of the effect of AAV on Ad replication and gene expression during coinfection revealed that the inhibitory effects of Rep are dependent both on the ratio of AAV to Ad, and on the particular time-point of infection (Timpe et al., 2006). In the presence of low concentrations of AAV (1 IU), E1A expression was in fact enhanced when compared to cells infected with Ad alone. Increasing levels of AAV resulted in repression of all Ad early promoters. This effect was not detectable until 12 h post infection however, correlating with the onset of Ad DNA replication. Indeed, Ad transcript levels remained unchanged prior to DNA replication even in the presence of overexpressed Rep78 (Timpe et al.,

2006). These results suggest that a combination of AAV amplification and Rep expression are necessary to inhibit Ad gene expression and DNA synthesis (Timpe et al., 2006), providing a potential mechanism for how AAV is able to repress Ad only after having benefited from the effects of its helper functions.

Independently from its inhibitory effects on Ad gene expression, AAV Rep also interferes with Ad by inhibiting Ad-mediated cellular transformation. One of the hallmarks of cellular transformation by adenovirus infection is a progression through S-phase triggered by the E1A-mediated release of the transcription factor E2F1 from pRb complexes, leading to cell proliferation. Indeed, this is a mechanism common to several other DNA tumor viruses, such as SV40 and HPV (Chellappan et al., 1992; Paganon et al., 1992). Ad also requires E2F1 for the transactivation of E2 via interactions with E4. Rep interferes with these pathways by (1) binding the E2F1 protein and stabilizing the pRb-E2F1 complexes, and, (2) binding the *E2F1* promoter directly and mitigating adenovirus-mediated *E2F1* transcription (Batchu et al., 2001). This serves to freeze cells in S phase and inhibit cell proliferation, and provides a molecular mechanism for the observed anti-oncogenic properties of AAV. Rep also independently achieves S-phase arrest through the binding and subsequent inhibition of Cdc25.

Similar observations have been made for HSV-1 supported AAV replication (Bantel-Schaal and Zur Hausen, 1988; Glauser et al., 2007, 2010). In the case of HSV-1, helper factors are all expressed with immediate early (IE) or early kinetics preceding viral replication, and are involved mostly in AAV DNA replication. Live cell covisualization experiments revealed the formation of separate AAV and HSV-1 RCs, with recruitment of Rep only in AAV RCs and HSV-1 ICP8 in both AAV and HSV-1 RCs (Glauser et al., 2007). Formation of HSV-1 RCs was dramatically reduced in the presence of AAV, as well as in the presence of Rep alone. Western blotting showed that the HSV-1 early proteins ICP4 and ICP8, the latter of which is a helper factor for AAV replication, were only modestly affected, while late proteins VP16 and gC were strongly reduced in the presence of AAV (Glauser et al., 2007). Although the mechanism of repression is not yet understood, it has been suggested that activation of the cellular DDR may in part be responsible.

Inhibition of HSV-1 is dependent on the ATPase and DNA binding activities of the Rep proteins (Glauser et al., 2010), which were also shown necessary to trigger DNA damage (Glauser et al., 2010; Schmidt et al., 2000). In addition, AAV inhibition of HSV-1 is associated with an increase in phosphorylation of the DNA damage marker, RPA, which has not been reported in the context of HSV-1 infection alone (Glauser et al., 2010).

Although there is still much to learn about the complex triangle of interaction between AAV, host, and helper virus, these studies provide crucial insight into the mechanisms through which a genomically simple virus such as AAV can navigate the complexity of interactions between itself, its host, and its helper virus. The requirement of AAV for a helper virus inevitably leads to competition for cellular resources as well as for the helper factors themselves, which are often essential for both AAV and the helper virus. Understanding how AAV manages to control its environment to its benefit may thus provide crucial insights into AAV biology.

1.4 THE COREPRESSOR KAP1

1.4.1 KAP1 overview

KAP1 (KRIP1/TIF1β/TRIM28) was first identified as an interaction partner for members of the Kruppel-associated Box (KRAB)-domain containing ZNF transcription factors (KRAB-ZFP) in several independent studies during the mid-1960s (Le Douarin et al., 1996; Friedman et al., 1996; Kim et al., 1996; Moosmann et al., 1996). KAP1 is a member of the tripartite motif (TRIM) family of proteins, many of which have previously been implicated in antiviral immunity (Ozato et al., 2008), and is very closely related to three other members of this family, TIF1α, TIF1γ, and TIF1δ, which have little functional overlap despite sharing numerous structural features (Peng et al., 2002). KAP1 is ubiquitously expressed during development and is thought to be a critical regulator of normal development and differentiation (Cammas et al., 2000, 2002, 2004; Jakobsson et al., 2008). Indeed, *Kap1* deletion in mice leads to embryonic lethality (Cammas et al., 2000). Additionally, KAP1 is involved in maintaining pluripotency and has also been linked to the proliferation and differentiation of tumor cells (Beer et al., 2002;

Seki et al., 2010; Silva et al., 2006; Xu et al., 2015). The pleiotropic effects of KAP1 are thought to occur through its capacity to regulate the dynamic organization of higher order chromatin structure through epigenetic modifications. KAP1 is recruited to the genome through interactions between its RBCC domain and KRAB-ZFP, whereupon it serves as a scaffold for the recruitment of numerous histone and nucleosome modifying proteins, which act together to generate transcriptionally repressive heterochromatin (Sripathy et al., 2006).

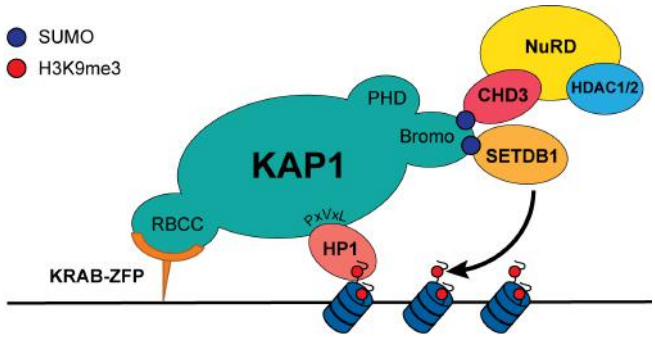


Figure 10. KAP1 protein interactions and mechanistic overview
KAP1 leads to the formation of transcriptionally repressive heterochromatin through the recruitment of various histone and nucleosome remodeling proteins. KAP1 is recruited to the genome through interactions between its RBCC protein interaction domain and the KRAB domain of KRAB-ZFP. Upon auto-SUMOylation of its bromodomain, KAP1 then serves as a scaffold for the recruitment of the histone deacetylase CHD3, a component of the NuRD complex, and the histone methyltransferase SETDB1, leading to the deposition of H3K9me3 and forming high affinity binding sites for HP1 bound to the central KAP1 domain.

1.4.2 Protein structure and interactions of KAP1

KAP1 shares the common overall architecture of the TIF1 subfamily of TRIM proteins, which consists of an N-terminal tripartite motif (TRIM) – containing a RING (really interesting new gene) finger, two B-box zinc fingers, and a coiled-coil domain (RBCC) – a central TIF1 signature sequence (TSS), and a C-terminal plant homeodomain (PHD) and bromodomain (Friedman et al., 1996; Venturini et al., 1999). In addition, KAP1 contains a central heterochromatin protein 1 (HP1) binding domain

also present in TIF1α and TIF1δ, which is essential for transcriptional repression by KAP1 (Figure 11).

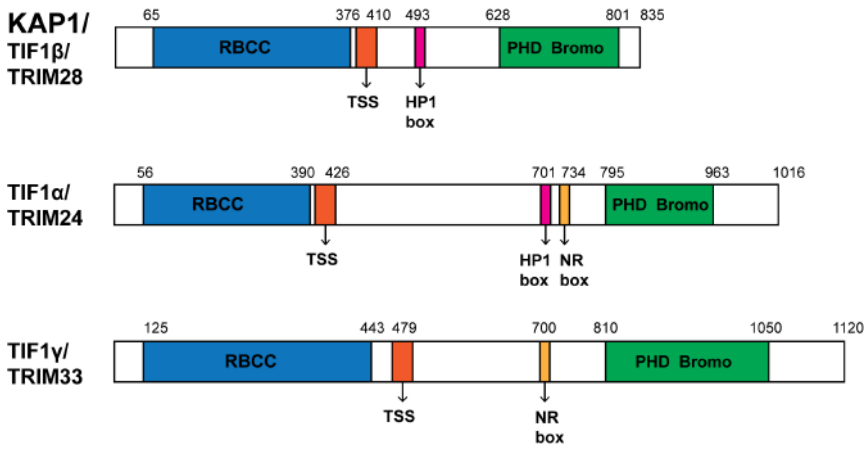


Figure 11. KAP1 domains
Schematic representation of KAP1 (also called TIF1β and TRIM28) and the related TIF1 family proteins, TIF1α/TRIM24 and TIF1γ/TRIM33. The RBCC domain, necessary for interactions between KAP1 and KRAB-ZFP, is shown in blue; the TSS site is shown in orange; the HP1 binding site is shown in pink; an NR box shown to bind nuclear receptors is shown in yellow; the C-terminal PHD and bromodomain (PB domain) are shown in green.

The N-terminal RBCC domain (aa 65-376) of KAP1 is a high-affinity protein interaction domain found in an increasing number of proteins with important roles in cell differentiation, development, oncogenesis, and apoptosis, which is necessary and sufficient to mediate interactions between KAP1 and the KRAB repression module of KRAB-ZFPs, and is thus responsible for the site-specific genomic recruitment of KAP1 (Borden, 1998; Peng et al., 2002; Saurin et al., 1996). In the context of KAP1, the RBCC domain homo-oligomerizes to bind a single KRAB module as a trimer (Peng et al., 2000). While all TIF1 family members contain an RBCC domain, KRAB-binding is specific to KAP1 (Peng et al., 2002). The defining element of the RBCC domain is the RING finger, which is composed of two consensus C3HC4 zinc binding motifs, and is found in more than 200 proteins that form macromolecular complexes with diverse functions (Borden, 1998; Saurin et al., 1996). The two B-boxes adjacent to the RING domain are cysteine + histidine-rich regions that also contain CHC3H2 zinc binding

motifs. The third sequence is the coiled-coil motif, an amphipathic double α-helix structure that, together with the two B-boxes, provides an extended hydrophobic region necessary to mediate protein interactions by the RBCC domain (Peng et al., 2000).

Adjacent to the RBCC domain is the TSS motif, a highly conserved 25 aa sequence rich in tryptophan and phenylalanine that is specific to the TIF1 subfamily of proteins. The role of the TSS is not yet clear, however its deletion in TIF1γ results in a complete loss of repressor activity (Venturini et al., 1999). The central domain of KAP1 is rich in prolines and glycines, and as such it is difficult to make predictions about the secondary structure of this region (Lechner et al., 2000). The observation that virtually the entire central domain of KAP1 is susceptible to protease treatment and that it has little α-helix or β-strand structure suggests that it exists in a highly extended, and perhaps flexible, conformation (Lechner et al., 2000), which has been suggested to provide KAP1 the adaptability necessary to interact with numerous different protein complexes (Iyengar and Farnham, 2011). While this central region is the least conserved domain among members of the TIF1 family and is in fact absent in TIF1γ, studies using recombinant GST-tagged KAP1 containing various deletions between aa 478-516 were able to map the KAP1 HP1 binding domain (HP1BD) to a 5 aa stretch within the central domain, comprised of a conserved PxVxL HP1 binding motif (Lechner et al., 2000). HP1 proteins bind trimethylated histone 3 lysine 9 (H3K9me3) and are essential for transcriptional repression by KAP1 (Lechner et al., 2000; Nielsen et al., 1999; Ryan et al., 1999). Each KAP1 HP1BD interacts with two HP1 monomers, and so a functional trimerized KAP1 corepressor might bring together six or more HP1 molecules (Lechner et al., 2000). Additionally, TIF1α and TIF1γ contain an NR box shown to bind nuclear receptors not present in KAP1.

The C-terminal tandem PHD and bromodomain (PB domain) of KAP1 (aa 628-801) form a cooperative unit that is required for transcriptional repression through the recruitment of chromatin modifying histone deacetylases and methyltransferases. Substitution of highly related PHD or bromodomains failed to restore KAP1 repressor activity, indicating a high specificity to their cooperative function and also demonstrating that both

domains are necessary for optimal repression (Schultz et al., 2001). The PHD finger is a 60 aa domain consisting of two C4HC3 zinc binding motifs arranged between anti-parallel β sheets, while the bromodomain is a 100 aa stretch consisting of a four-helical bundle (Capili et al., 2001) whose conserved hydrophobic core recognizes the backbone of histone tails (Schultz et al., 2001). The PHD domain of KAP1 functions as an intramolecular E3 ligase, auto-sumoylating the adjacent bromodomain (Ivanov et al., 2007), after which the sumoylated bromodomain can then recruit two chromatin-modifying enzymes via their SUMO interaction motifs (SIM) – (1) chromodomain-helicase-DNA-binding protein 3 (CHD3/Mi2α) found in the Nucleosome Remodeling and Histone Deacetylation (NuRD) complex, and (2) SET domain, bifurcated 1 (SETDB1), a histone methyltransferase specific for H3K9 (Schultz et al., 2001, 2002; Zeng et al., 2008). Trimethylation of H3K9 by SETDB1 subsequently creates high affinity binding sites for HP1, stabilizing the interaction between heterochromatin and the KAP1 macromolecular complex.

1.4.3 Recruitment of KAP1 to the genome

Lacking any intrinsic DNA binding activity, KAP1 is dependent on interactions with KRAB-ZFP for recruitment to the genome. Numerous studies had demonstrated that the fusion of KRAB domains to heterologous DNA binding domains resulted in transcriptional repression of target sequence-containing promoters in a dose-dependent manner and over long distances (Margolin et al., 1994; Witzgall et al., 1993, 1994). KRAB domains were first identified as a conserved sequence at the N-terminal end of proteins containing C2H2 zinc fingers. The C2H2 family of zinc fingers represents the largest class of DNA-binding transcription factors in the human genome, and it is now clear that approximately half of these contain KRAB domains (KRAB-ZFP) (Urrutia, 2003). There are ~742 structurally different KRAB-ZFP that have evolved through duplication events to perform species- and tissue-specific gene regulation, and together these proteins regulate essential processes during development and differentiation (Emerson and Thomas, 2009; Urrutia, 2003). Mammalian KRAB-ZFP contain up to 30 N-terminal ZNF domains (Emerson and Thomas, 2009) each of

which can recognize 3 nt of DNA. In this way, KRAB-ZFPs become anchored to specific target DNA sequences where they can then recruit KAP1 and, consequently, the NuRD complex, SETDB1, and HP1.

Site-directed mutagenesis of KRAB domains demonstrated that these regions were responsible for the observed transcriptional repression of reporter constructs (Witzgall et al., 1994). Shortly thereafter, yeast two hybrid experiments identified KAP1 as the binding partner for two KRAB-ZFPs, KOX-1 and KIP-1 (Kim et al., 1996; Moosmann et al., 1996), and direct tethering of KAP1 to DNA demonstrated that KAP1 was in fact the mediator of the repression exerted by these KRAB-ZFP (Friedman et al., 1996; Moosmann et al., 1996). KAP1 has since been implicated as a general corepressor of three distinct subfamilies of KRAB-ZFP (Abrink et al., 2001), and indeed, 51 of 61 tested human KRAB-ZFPs interact with and mediate repression through KAP1 (Itokawa et al., 2009). Additionally, certain KRAB proteins in which alternative splicing has removed the C-terminal C2H2 zinc finger motifs have been shown to serve as adaptor molecules for KAP1-mediated repression. The KRAB only (KRAB-O) protein, a splice variant of *Zfp748,* tethers KAP1 to the sex determination transcription factor SRY, recruiting KAP1 to SRY binding sites (Peng et al., 2009). Similarly, the von Hippel-Lindau tumor suppressor (pVHL)-associated KRAB-A domain containing protein (VHLaK) is a splice variant of *Zfp197* that recruits KAP1 to pVHL, resulting in the pVHL-mediated transcriptional repression of hypoxia-inducible factor-1α (HIF-1α) (Li et al., 2003)

To determine the *in vivo* relevance of KAP1 interactions with KRAB-ZFPs and other DNA-binding proteins, chromatin-immunoprecipation (ChIP) experiments were used to identify genomic KAP1 binding sites. Complete genomic analysis of Ntera human embryonic carcinoma cells identified over 7000 KAP1 binding sites, mapping either to the 3' coding regions of ZNF genes, or core promoter regions near transcription start sites (TSS) (O'Geen et al., 2007). The strongest KAP1 binding sites were at the 3' coding regions of ZNF genes however, leading the authors to suggest that KRAB-ZFPs participate in an autoregulatory loop involving recruitment of KAP1. ChIP experiments using various KAP1 mutant proteins demonstrated that recruitment of KAP1 to the 3' coding regions of ZNF genes was in fact

dependent on the KAP1 RBCC domain (Iyengar et al., 2011), and further studies showed that ZNF274 colocalizes with KAP1 at these regions (Frietze et al., 2010), supporting the notion that KAP1 is recruited to the genome via interactions with KRAB-ZFPs. Interestingly however, binding of KAP1 to TSS sites near promoter regions was independent of the KAP1 RBCC domain, indicating a second mechanism for KAP1 recruitment to these sites not involving KRAB-ZFP (Iyengar et al., 2011). Although the exact mechanism for this recruitment is still unclear, mutational analysis of KAP1 suggests that protein-protein interactions with aa 380-618 within the KAP1 central domain may be involved (Iyengar et al., 2011).

1.4.4 Role of KAP1 in transcriptional regulation

Several lines of evidence suggest that KAP1 functions as a transcriptional repressor by coordinating the activities of nucleosome- and histone-modifying proteins to induce histone methylation, HP1 deposition, and the formation of transcriptionally repressive heterochromatin. The majority of data demonstrating repression by KAP1 however come from artificial recruitment experiments in which isolated KRAB domains fused to heterologous DNA binding domains were shown to mediate robust repression of reporter constructs containing target DNA sequences in a manner dependent on KAP1 binding (Margolin et al., 1994; Witzgall et al., 1994). Several studies using inducible KRAB fusion proteins were able to demonstrate that KAP1-mediated repression of integrated reporter constructs was associated with a recruitment of SETDB1 and the HP1 family of proteins, as well as the deposition of the repressive mark, H3K9me3 (Ayyanathan et al., 2003; Schultz et al., 2002; Sripathy et al., 2006). A hormone inducible KOX1 KRAB domain fused to the PAX3 DNA binding domain was shown to repress transcription from an integrated TK-luciferase reporter construct containing 6 high affinity PAX3 binding sites upstream of the TK promoter. Importantly, a mutant KRAB domain unable to bind KAP1 had no effect (Ayyanathan et al., 2003). Using immunofluorescence and ChIP experiments, the investigator further demonstrated that repressed reporter constructs were spatially recruited to HP1-rich nuclear regions, and that they physically associated with all components of the KAP1 repressive

complex – KAP1, SETDB1, HP1α/γ, and H3K9me3 (Ayyanathan et al., 2003). Furthermore, repression resulting from one short exposure to the hormone inducer was heritable for up to 50 generations (Ayyanathan et al., 2003). Similar results were obtained using hormone inducible GAL4-KRAB fusions and an integrated TK-luciferase reporter containing 5 upstream GAL4 binding sites (Sripathy et al., 2006). Additionally, siRNA-mediated depletion of endogenous KAP1, HP1α/β/γ, and SETDB1 attenuated repression of the reporter, and direct tethering of WT KAP1, but not KAP1 with mutations in the HP1 binding site or PB domain, was sufficient to repress transcription from the integrated transgene (Sripathy et al., 2006).

Genomic KAP1 binding sites at the 3' coding regions of ZNF genes have also been shown to be enriched for SETDB1, HP1, and H3K9me3 (Frietze et al., 2010; Iyengar et al., 2011; Vogel et al., 2006). A systematic investigation of the target genes of HP1β revealed a significant enrichment for this heterochromatin-inducing protein at these sites (Vogel et al., 2006), and sequential ChIP analysis of K562 cells further demonstrated colocalization of SETDB1, KAP1, and H3K9me3 deposition (Frietze et al., 2010). While these characteristics have clearly been shown to mediate repression in the context of artificial recruitment experiments, the *in vivo* functional relevance of KAP1-SETDB1 recruitment to ZNF genes is less clear. Depletion of KAP1 by siRNA had only minor effects on the transcriptome of HEK293 cells (Iyengar et al., 2011). Furthermore, the set of ~90 genes that were responsive were not those bound at ZNF 3' exons, but instead showed KAP1 binding only at distant intergenic sites (Iyengar et al., 2011), suggesting either an indirect route to regulation or long-distance heterochromatin spreading (Groner et al., 2010). It has been suggested that, rather than playing a role in transcriptional regulation, KAP1 binding to ZNF 3' exons may play a role in stabilizing KRAB-ZNF gene clusters (Vogel et al., 2006). KRAB-ZFP have undergone a rapid expansion in recent evolution, and sequence comparisons of various KRAB-ZFP gene clusters indicate that in situ tandem duplications play an important role in the evolution of these genes (Dehal et al., 2001). It is thought that the heterochromatinization of these regions may prevent the recombination-mediated deletion of newly

duplicated KRAB-ZFP, and may in fact have facilitated the incredible expansion of this gene family (Vogel et al., 2006).

In contrast to KAP1 binding sites at 3' ZNF exons, binding sites near TSS are neither dependent on KRAB-ZFP for recruitment of KAP1, nor associated with enrichment for H3K9me3 (Frietze et al., 2010; Iyengar et al., 2011). Rather, these binding sites may be related to the recently discovered role for KAP1 in promoter proximal pausing via recruitment of the 7SK small nuclear ribonucleoprotein (snRNP) (Bunch et al., 2015; McNamara et al., 2016). Promoter proximal pausing of RNA polymerase II (RNA pol II) represents a major checkpoint for transcription. Transcription pre-initiation complexes (PIC) assemble at promoters to recruit RNA pol II and initiate transcription (Diamant and Dikstein, 2013; Hargreaves et al., 2009; Smale, 2010), however RNA Pol II pauses ~20-63 nt downstream of TSS through the activity of various negative elongation factors (Adelman and Lis, 2012; Peterlin and Price, 2006; Yamaguchi et al., 1999). For transcription to continue, RNA pol II must be phosphorylated at its C-terminal domain (CTD) by a positive elongation factor, such as the P-TEFb kinase (Mancebo et al., 1997; Peng et al., 1998). To prevent premature phosphorylation of RNA pol II, P-TEFb is held in a catalytically inactive state, reversibly bound to the 7SK snRNP complex (He et al., 2008; Jeronimo et al., 2007; Krueger et al., 2008).

The role for KAP1 in this mechanism was first identified through an unbiased search for human proteins involved in regulating RNA pol II pausing using *HSP70* as a model gene (Bunch et al., 2014). Genome wide ChIP-sequencing (ChIP-seq) analysis of WT and KAP1 depleted cells revealed a strong correlation between KAP1 and RNA pol II occupancy. Furthermore, the investigators showed that KAP1 rapidly became phosphorylated at serine 824 (a modification important in DNA damage repair, reviewed in section 1.4.5) in a DNA-PK and ATM dependent manner upon heat shock, and that this correlated with RNA Pol II progression and transcriptional elongation of *HSP70* (Bunch et al., 2014). This mechanism was further investigated in the regulation of transcription initiation from primary response genes (PRGs) (McNamara et al., 2016), which utilize pTEFb to rapidly phosphorylate paused RNA pol II in response to danger stimuli. KAP1 was shown to bind directly to the 7SK snRNP complex via

interactions with the Larp7 component. The investigators comprehensively demonstrated that KAP1 is dynamically recruited to the promoters of PRGs, both prior to stimulation and throughout active transcription, where it tethers the 7SK snRNP complex and its associated pTEFb (McNamara et al., 2016). This functions to provide a constant source of pTEFb ready to phosphorylate and activate RNA pol II upon stimulation. Supporting this, KAP1 depletion resulted in a sharp reduction of 7SK snRNP occupancy and pTEFb recruitment at PRG promoters, which interestingly correlated with a decrease in magnitude of gene induction upon stimulation, suggesting that KAP1 is necessary for proper gene induction of PRGs. Furthermore, genome-wide studies revealed that KAP1 and 7SK snRNP co-occupy most promoter-proximal regions containing paused RNA Pol II, also establishing a role for KAP1 in maintenance of efficient levels of basal transcription.

1.4.5 Role of KAP1 in the cellular DDR

Transcriptional repression by KAP1 is dependent on the attachment of small ubiquitin-like modifier (SUMO) proteins to its bromodomain, which function to recruit SETDB1 and the NuRD complex protein CHD3 via their SUMO-interacting motifs (SIM). SUMO family proteins are conjugated to target lysines via a cascade of E1, E2, or E3 ligases (Isogai and Shirakawa, 2007), where they appear to play a key role as post-translational modifications involved in transcriptional repression (Gill, 2005). The KAP1 PHD domain is highly related to the RING finger, often found in ubiquitin and SUMO E3 ligases (Joazeiro and Weissman, 2000), and indeed, the PHD domain acts as an intramolecular E3 SUMO ligase to conjugate SUMO moieties to the adjacent bromodomain, predominantly at lysines, 554, 770, and 804 (Ivanov et al., 2007; Lee et al., 2007). This activity was shown to be essential for KAP1 recruitment of CHD3 and SETDB1, and consequently for the repressor activity of KAP1 (Ivanov et al., 2007; Lee et al., 2007; Mascle et al., 2007)).

KAP1 SUMOylation and repressor activity is in turn regulated through phosphorylation of its C-terminal domain. In response to DNA damage, the kinases ATM and DNA-PK mediate phosphorylation of KAP1 at serine 824 (p-KAP1-S824), adjacent to the SUMOylated bromodomain (Tomimatsu et

al., 2009; White et al., 2006; Ziv et al., 2006). This has been shown to lead to the direct upregulation of the KAP1-regulated genes *Gadd45α* and *p21,* involved in cell cycle arrest, as well as to modulate expression of various other KAP1-targeted genes involved in the cellular DDR, including the proapoptotic genes *Bax, Puma,* and *Noxa* (Li et al., 2007b). KAP1 also has a more structural role in DNA repair. ATM activation upon the induction of DSBs leads to the phosphorylation of KAP1-S824 at sites of damage, which then spreads throughout the chromatin. This results in a wave of global chromatin decondensation that has been shown to increase survival after ionizing radiation, and, more importantly, to be essential for the repair of DSBs within heterochromatin by increasing access to essential repair factors (Goodarzi et al., 2008; Noon et al., 2010; Ziv et al., 2006). Phosphorylated KAP1-S824 colocalizes at damage foci with numerous other proteins necessary for DNA repair, such as γH2AX, 53BP1, and ATM, and studies using non-phosphorylatable KAP1 (KAP1-S824A) demonstrated a constitutive heterochromatic DSB repair defect characterized by a loss of chromatin decondensation, supporting a direct role for KAP1 in this process (Noon et al., 2010; Ziv et al., 2006).

The exact mechanism through which phosphorylation of KAP1 leads to the relaxation of chromatin and efficient DNA repair is not yet clear. Studies showing that the induction of DNA damage is associated with a decrease in levels of SUMOylated KAP1 suggest that phosphorylation of KAP1-S824 negatively regulates SUMOylation levels (Li et al., 2007b), which would in turn lead to reduced interactions between KAP1, SETDB1, and CHD3, and thus chromatin relaxation. Supporting this, the deSUMOylase SUMO/Sentrin/Smt3-specific peptidase (SENP70) interacts with KAP1 via HP1α to promote the removal of SUMO from KAP1, and this was shown to regulate the interaction between KAP1 and CHD3 (Garvin et al., 2013). Indeed, CHD3 dispersion from sites of DSBs upon phosphorylation of KAP1-S824 was associated with chromatin relaxation (Goodarzi et al., 2011). However, the investigators in this study did not find any change in SUMOylation of KAP1 between phosphorylated and non-phosphorylated forms, and instead suggested that phosphorylation at S824 interferes with

the binding between SUMO bound to KAP1 and the SIM of CHD3 through charge repulsion (Goodarzi et al., 2011).

KAP1 can also become phosphorylated at serine 473 (p-KAP1-S473) either by the ATM-Chk2 or ATR-Chk1 pathway depending on the type of DNA damage that occurs (Chang et al., 2008; Hu et al., 2012). Differences in the kinetics and nuclear localization of p-KAP1-S824 and p-KAP1-S473 suggest that each phosphorylated form of KAP1 participates in DNA repair in spatially and temporally regulated way, acting synergistically to repair damaged DNA (Hu et al., 2012). Whereas phosphorylation of KAP1-S824 interferes with interactions between KAP1, SETDB1, and CHD3, the proximity of S473 to the PxVxL motif of the HP1 binding domain leads to a loss of interaction between p-KAP1-S473 and HP1 (Chang et al., 2008; Hu et al., 2012). Studies using a phosphomimetic mutant of KAP1-S473 demonstrated that the loss of HP1 binding resulted in the upregulation of various KAP1-regulated genes, including *p53AIPl, Noxa, Zfp180,* and *RNApol II*. Loss of HP1 binding was also associated with increased binding of KAP1 to the transcription factor E2F1, thus reducing the ability of E2F1 to induce expression of E2F1-regulated proapoptotic genes (Hu et al., 2012). Additionally, phosphorylation of KAP1-S473 by protein kinase C-δ (PKCδ) during early S phase was shown to dynamically regulate cell cycle progression through the induction of cell cycle genes *cyclin A2, Cdc2,* and *Cdc25,* leading to G_2/M progression (Chang et al., 2008; Lee et al., 2012).

Regardless of the site of KAP1 phosphorylation, basal homeostatic levels of phosphorylated/SUMOylated KAP1 must be restored upon completion of DNA repair or cell cycle progression in order to restore steady state gene transcription and heterochromatin formation. Protein phosphatase 4 (PP4) was shown to interact with KAP1, and depletion of PP4 resulted in prolonged phosphorylation of both S824 and S473, resulting in prolonged G_2/M checkpoint and relaxation of heterochromatin with an associated release of CHD3 (Lee et al., 2012). Two isoforms of PP1 were also shown to dephosphorylate KAP1-S824 under different stress conditions (Li et al., 2010). PP1α interacted with KAP1 under steady state conditions, while PP1β interacted with KAP1 under conditions of genotoxic stress. However, based on experiments using the PP1 inhibitory peptide I-2, the investigators

concluded that PP1α could be one of the key molecules to regulate levels of damage-induced p-KAP1-S824. Depletion of PP1α had a similar effect on the expression of KAP1-regulated genes as was observed with depletion of KAP1, and it was subsequently found to form a constitutive functional unit with KAP1 at the *p21* proximal promoter (Li et al., 2010).

Finally, KAP1 has also been shown to inhibit apoptosis through the modulation of post-translational modifications of various transcription factors independently of DNA-binding. Mouse double minute 2 homolog (MDM2) is an E3 ubiquitin ligase that functions to promote p53 ubiquitination and degradation, thus playing an essential role in the regulation of apoptosis in response to mitogenic signals. KAP1 binds MDM2 via its coiled-coil domain, forming a ternary complex with p53 and cooperating with MDM2 to inhibit p53 acetylation through the tethering of KAP1-associated histone deacetylases (HDAC) and thus inhibiting p53 transcriptional and apoptosis functions (Wang et al., 2005). KAP1 also binds directly to the transcription factor E2F1 via its coiled-coil domain, stimulating the formation of E2F1-HDAC1 complexes that inhibit E2F1 acteylation/stimulation, and preventing E2F1-mediated apoptosis (Wang et al., 2007). Finally, KAP1 was also found to act as a transcriptional regulator of the signal transducer and activator of the transcription 3/interleukin-6 (STAT3/IL-6) pathway by associating with STAT3, an interaction which prevents the accumulation of phosphorylated STAT3 (Tsuruma et al., 2008), although the mechanism behind this observation is still unclear.

1.4.6 Role of KAP1 in the regulation of viral elements

The role for KAP1 in the regulation of viral elements was first discovered by Wolf and Goff (Wolf and Goff, 2007) who showed that recruitment of KAP1 by the mouse-specific KRAB-ZFP ZFP-809 was responsible for the restriction of murine leukemia virus (MLV) replication in mouse embryonic stem (ES) and embryonic carcinoma (EC) cells (Wolf and Goff, 2007; Wolf et al., 2008). It was known that MLV could successfully infect and integrate into the genomes of these cells, but that potent repression of the viral promoter in the long terminal repeat (LTR) silenced these genomes through a mechanism involving recruitment of a *trans* acting

DNA-binding factor (Akgün et al., 1991; Flanagan et al., 1989; Loh et al., 1988; Tsukiyama et al., 1989) to the viral repressor binding site within the viral primer binding site (PBS) (Barklis et al., 1986; Feuer et al., 1989; Loh et al., 1987), resulting in methylation of the provirus (Niwa et al., 1983). KAP1 was successfully purified from, and was demonstrated to be an integral component of, the repressive complex bound to the viral repressor binding site (Wolf and Goff, 2007). Reactivation of silenced MLV in differentiated EC cells was associated with a decrease in steady-state levels of KAP1, and KAP1 depletion in two EC cell lines was shown to reduce repression of an integrated MLV-based retroviral vector containing the WT MLV PBS. ChIP experiments confirmed that KAP1 was bound to MLV-based vectors containing a WT PBS, which are repressed, but not to vectors containing a mutant PBS (B2), which are not repressed (Wolf and Goff, 2007). Silencing of the MLV via the PBS was further shown to be dependent on HP1 recruitment by KAP1, further supporting its role in MLV restriction (Wolf et al., 2008). Based on this, it was suggested that KAP1 has an essential role in orchestrating PBS-mediated silencing of integrated MLV proviruses in EC and ES cells, and furthermore that this may represent a strategy evolved to protect the embryo from the reactivation of retrotransposons and endogenous retroviruses during embryogenesis (Wolf and Goff, 2007).

Transposable elements (TE) account for more than half of the human and murine genomes (Lander et al., 2001; Waterston et al., 2002), and their ability to mediate recombination events, retrotranspose, and alter the transcription of neighboring genes through strong promoter/enhancer elements poses a threat to the developing embryo, particularly after the genome-wide DNA demethylation that occurs several times during embryogenesis (Morgan et al., 2005). It is now known that KAP1-mediated repression of various TE, in particular endogenous retroviruses (ERV), is crucial for the maintenance of genomic integrity during embryogenesis, and for the preservation of transcriptional dynamics, both during early development and in certain differentiated cells (Ecco et al., 2016; Rowe and Trono, 2011; Rowe et al., 2010, 2013). Depletion of KAP1 led to a marked increase in expression of ERVs, in particular intracisternal A-particles (IAP), in mouse embryonic carcinoma (EC) and embryonic stem (ES) cells, and

KAP1 docking at these sites triggered the deposition of repressive marks such as H3K9me3 and DNA methylation (Rowe and Trono, 2011; Rowe et al., 2010). Further studies demonstrated that KAP1- and SETDB1-mediated histone methylation acted cooperatively to induce *de novo* DNA methylation of flanking regions (Rowe et al., 2013), possibly through the direct recruitment of DNA methylases by KAP1 and SETDB1 (Li et al., 2006; Quenneville et al., 2011), leading to the silencing of strong neighboring promoters (Rowe and Trono, 2011; Rowe et al., 2013). The result was the establishment of stable site-specific DNA methylation signatures across the genome in early embryogenesis, presumably necessary to prevent retrotransposition as well as the illegitimate expression of neighboring genes. The mouse-specific KRAB-ZFP ZFP932 and its paralog Gm15446 were later shown to bind with KAP1 to numerous ERV (226 for ZFP932 and 448 for Gm15446) in mouse ES cells, and knockout of the KRAB-ZFPs led not only to the upregulation of their target ERV (amongst others), but also to the upregulation of 71 neighboring genes (Ecco et al., 2016). The investigators found that, although less pronounced, KAP1 knockout in certain differentiated mouse cells also led to an upregulation of different subsets of ERVs and associated neighboring genes. This effect was not systematically induced in all cell lines tested however, and it was suggested that the presence or absence of tissue-specific activators is necessary for the upregulation of KAP1-regulated gene expression upon ablation of repressive methylation (Ecco, 2016).

More recently, KAP1 has been shown to be involved in the repression of human cytomegalovirus (CMV) (Rauwel et al., 2015) and Kaposi's sarcoma-associated herpesvirus (KSHV) (Chang et al., 2009; Sun et al., 2014; Zhang et al., 2014). In the context of these latent DNA viruses however, KAP1 appears to act more as a regulator of latency. During KSHV infection, KAP1 is recruited to the genome by the viral latency associated nuclear antigen (LANA) protein, where its repression of lytic genes, such as k-bZIP and vPK, is crucial to the maintenance of latency (Sun et al., 2014). Induction of lytic replication through overexpression of the viral transcriptional regulator K-Rta results in the disassociation of KAP1 and the subsequent loss of HP1 and H3K9me3, and, consequently, reactivation of

the viral genome (Chang et al., 2009). As reactivation progresses, the newly expressed viral kinase vPK inactivates KAP1 corepressor activity through phosphorylation of KAP1 serine 824, suggesting counteraction by the virus is necessary for full reactivation and maintenance of lytic replication (Chang et al., 2009). A more recent study has also shown KAP1 to be necessary for both the establishment and maintenance of HCMV latency in hematopoietic stem cells (HSCs) (Rauwel et al., 2015). Here again, latency was shown to correlate with KAP1-mediated recruitment of SETDB1 and HP1, and with H3K9me3 deposition across the viral genome. Accordingly, depletion of KAP1 in HSCs resulted in the upregulation of immediate early, early, and late viral genes. Importantly, binding of SETDB1 and H3K9me3 were lost upon differentiation of HSCs to permissive dendritic cells (DCs) (Rauwel et al., 2015).

Finally, in a mechanism similar to the manner in which KAP1 modulates the post-translational modifications of certain TFs, KAP1 has been shown to inhibit integration of HIV by forming a complex with the viral integrase (IN) causing its inactivation through deacetylation by KAP1-associated HDAC (Allouch et al., 2011). KAP1 was also suggested to repress adenoviral replication at early stages of infection by preventing efficient decondensation of incoming viral genomes, an effect which was counteracted by crosstalk between the Ad E1B55K protein and KAP1, resulting in decreased SUMOylation and increased phosphorylation of KAP1 (Bürck et al., 2015; Schreiner et al., 2013a).

1.5 AIMS OF THE THESIS

As outlined above, numerous interactions between AAV Rep and cellular proteins are necessary to support the complex biphasic life cycle of AAV. While a multitude of cellular Rep interaction partners have previously been described, mechanistic insights into these interactions are largely missing. Additionally, several aspects of AAV biology remain elusive despite our current knowledge of Rep protein interactions, suggesting that further interactions remain to be identified. One particularly intriguing feature of the AAV Rep proteins that remains to be fully elucidated is their capacity to regulate transcription from both AAV and heterologous viral and cellular promoters via a mechanism that is likely dependent on interactions with cellular transcriptional machinery. With this in mind, the aims of this thesis were to (1) identify new cellular interaction partners for the Rep proteins, (2) identify their role in the AAV life cycle, focusing on Rep-mediated transcriptional regulation, and (3) determine their mechanism of action. It is important to acknowledge here that various other aspects of basic AAV biology are also still poorly understood. In particular, little is yet known about the nature and contribution of epigenetic marks to the genome organization and temporal gene regulation of AAV. It is therefore that we approached the downstream analysis of putative Rep interaction partners with an open mind. In addition to enhancing our general understanding of AAV, identifying essential Rep-protein interactions will have implications for AAV-based gene therapy, as significant enhancements in AAV vector design have previously been derived from an enhanced understanding of basic wild type AAV2 biology.

Described in Results Chapter 3 is the development and application of a screening method known as BioID, which was used to generate a list of potential interaction partners for the Rep proteins. Several candidates were chosen for downstream analysis based on cellular function and/or known interactions with viral elements, one of which was the transcriptional corepressor KAP1. Results Chapters 4 and 5 focus on the discovery of an intricate web of interactions between the Rep proteins, KAP1, the phosphatase PP1, and its negative regulator NIPP1. In Results chapter 4, we show that, while KAP1 is not involved in Rep-mediated transcriptional

regulation, depletion of KAP1 enhances AAV2 replication and transcription. We further demonstrate that KAP1 binds the latent AAV2 genome, leading to trimethylation of AAV2-associated H3K9. Results Chapter 5 focuses on defining a mechanism of PP1 antagonism through which the AAV2 Rep proteins counteract KAP1 repression during lytic replication. Finally, in Results Chapter 6, we show that infection by Ad5 and HSV-1 leads to depletion of KAP1 and suggest that this may represent an unknown helper function necessary to relieve AAV from its latent state.

Chapter 2: Materials and Methods

2.1 MOLECULAR CLONING

2.1.1 Standard Polymerase Chain Reaction (PCR)

This is the basic PCR reaction that was used for colony screening and other tests. The polymerase kit used was the standard goTaq® DNA polymerase (Promega). PCR reactions were carried out in a total volume of 25µl containing (final concentrations): 200µM dNTPs (mix of dATP, dCTP, dGTP and dTTP, from NEB), 1X Colorless GoTaq Reaction Buffer, 1.25U GoTaq DNA polymerase, 0.4µM forward and reverse primers, ~10ng DNA, and ddH_2O in a final volume of 25µl. 0.01µmol of unmodified oligonucleotide primers for PCR amplification were obtained from Eurofins MWG Operon and re-suspended with ddH_2O to make stock solutions at 100µM. Forward (sense) and reverse (anti-sense, reverse complement) oligonucleotide primers were designed following the manufacturer's directions: 15-30 nucleotides in length, 40- 60% GC content, and terminating in at least one G/C base. PCRs were performed using an Eppendorf Mastercycler EP Gradient Thermal Cycler. A typical PCR reaction initiated with a 2 minute denaturing step at 95℃ followed by 25-35 cycles including a denaturing step of 20 seconds at 95℃, an annealing step of 45 seconds at 50°C (depending on the primers' melting temperature) and an extension step of 1minute/kb at 72°C. A final 5 minutes extension at 72°C followed the cycles. Reactions were kept at 4°C until use. For information on all the PCR primers used in this thesis please see Table 1 in Appendix 1, pg. 239.

2.1.2 High fidelity PCR

The Phusion High Fidelity polymerase kit (Thermo Scientific) was used for all PCR reactions performed to amplify DNA fragments necessary for cloning. Where relevant, primers used for subcloning of DNA fragments into vectors were designed to include an optimal Kozak consensus sequence (GCC ACC) immediately upstream of the ATG, and a two-glycine linker (GGC GGC) between N-terminal FLAG or T7 tags and the start of the coding sequence. PCR reactions were carried out in a total volume of 25µl, in a

mixture containing (final concentrations) 10ng of DNA template, 0.5μM of forward and reverse primers, 200μM dNTPs (NEB), 0.5U/25μl Phusion® High-Fidelity DNA polymerase (Thermo Scientific), 1X Phusion® HF or GC Buffer and brought to the final volume with ddH_2O. PCRs were performed using an Eppendorf Mastercycler EP Gradient Thermal Cycler. Phusion GC buffer was used in reactions that initially did not work with HF buffer, as it is indicated for GC-rich templates or those prone to forming secondary structures. Reaction conditions were varied depending on the size and GC content of the fragment to be amplified. A typical PCR reaction initiated with a 30 seconds denaturing step at 98°C followed by 25-35 cycles including a denaturing step of 10 seconds at 98°C, an annealing step of 30 seconds at 68°C and an extension step of 30 seconds/kb at 72°C. A final 10 minutes extension at 72°C followed the cycles. Reactions were kept at 4°C until use. The annealing step temperature was adjusted to the melting temperature of the primers. Primers were designed and obtained from Eurofins MWG Operon as explained above. For some cloning strategies, primers were designed with external overhangs containing restriction sites matching those in the destination construct.

2.1.3 Overlapping PCR

Overlapping PCR was used to fuse BirA* to the N-terminus of the Rep proteins for BioID. BirA* was first amplified by high fidelity PCR using a forward primer (primer1) containing an optimal Kozak sequence and an *ApaI* restriction site, and a reverse primer (primer 2), in which the stop codon was replaced with a 10X glycine linker (as reverse complement: 5X GCC TCC repeats). Similarly, individual Rep fragments were amplified by high fidelity PCR using a forward primer (primer 3) containing the complementary glycine linker sequence for primer 2 (5X GGC GGA repeats) and a reverse primer (primer 4) containing an *ApaI* restriction site. Amplified products were then purified and used in a third PCR reaction using primer 1 and primer 4. For this reaction, primer concentration was reduced to 0.2μM, and 10ng of each fragment was used. Cycling conditions were as follows: an initial denaturation step of 98^{o}C, followed by 30 cycles including a denaturation step of 10 seconds, an annealing step of 15 seconds at 70^{o}C, and an

elongation step of 1 minute at 72^{o}C, followed by a final elongation step of 10 minutes at 72^{o}C.

2.1.4 Site-directed mutagenesis

To generate Rep point mutants the Stratagene site-directed mutagenesis kit was used. Primers were designed to contain the desired point mutations in the centre flanked by 10-15 base pairs on either side matching the template sequence to allow sufficient annealing to the template DNA. This technology was used to introduce the K372A mutation into Rep.

2.1.5 Agarose gel electrophoresis

Amplified DNA PCR products or DNA fragments from restriction enzyme (RE) digestions from PCR were diluted with 5X DNA loading dye (NEB). Agarose gels were prepared by dissolving agarose powder (UltraPure^{TM} Agarose, Life Technologies) in 1X TAE buffer (10X: Tris-Base 48.4g/L, Acetic acid 11.4ml/L, EDTA 3.7g/L) and heating until boiling. The agarose percentage varied between 0.7% and 1.5% depending on the size of the fragment to be observed/isolated, with higher percentages used to achieve better separation of smaller fragments. After allowing the dissolved agarose to cool down, 1µg/ml of Ethidium bromide solution (Sigma) was added and the mixture was poured into an electrophoresis tank (BioRad). Once solidified, the gel was immersed in 1X TAE buffer and DNA samples were loaded in the wells of the gel. Gels were run at 80-120V for about 30-90 minutes, depending on the separation required between bands. Band sizes were monitored with the 100bp or the 1kb DNA ladder (NEB), which was run in parallel. Ethidium bromide-stained DNA fragments were visualized under a ultra-violet trans-illuminator using a ChemiDocTM XRS+ System (BioRad).

2.1.6 Extraction and purification of DNA fragments

DNA fragments of interest were excised from agarose gels and purified using a QIAGEN Gel Extraction Kit. Three volumes of solubilisation and binding QG buffer were added to the agarose gel slice, the mixture was heated to 50°C for 10 minutes and vortexed to dissolve the slice. One

volume of isopropanol was added to the mixture and subsequently added to a QIAquick spin-column. DNA was bound to the column by centrifugation at 13,000 rpm for 1 minute, after which the flow-through was discarded and the column washed with 750µl of ethanol-containing PE buffer and centrifuged twice to completely remove the ethanol. The DNA was finally eluted into a sterile eppendorf tube with 30µl of ddH_2O. A similar procedure was used to purify RE digestion products that were not run on agarose gels. The same kit and protocol were used, with the exception that the RE digestion mixture was directly mixed with the QG buffer and isopropanol before proceeding with the column purification.

2.1.7 Restriction Enzyme Digestion of DNA

For analytical RE digestions, 2U of RE (NEB) was used to digest 500ng of plasmid DNA in a 1X solution of the appropriate buffer supplied by the manufacturer, and total reaction volume was adjusted to 20µl with ddH2O. For RE digestions to generate fragments for cloning, 5-10U of RE was used to digest 2-5µg of DNA. All reactions were performed at the indicated temperature for the used RE (generally at 37℃) for 2 h. When digestions with two restriction enzymes were performed, a compatible reaction buffer was used as indicated by manufacturer. After single-enzyme digestions to generate plasmid fragments for cloning, vector DNA was treated with calf intestinal phosphatase (CIP, NEB) to remove the 5' phosphate group from the digested vector to prevent re-circularisation between compatible ends. 5U of CIP were added directly to the RE digestion mixture and incubated at 37℃ for 30 minutes and further purified as described above.

2.1.8 DNA ligations

The purified DNA insert and vector were mixed at a 5:1 ratio using 50ng of vector with 1µl of 10X T4 ligase buffer (NEB) and 0.5µl of T4 ligase enzyme (NEB), in a total reaction volume adjusted to 10µl with ddH_2O. Reactions were incubated at room temperature for 2 hours or overnight at 16℃.

2.1.9 Competent bacterial cells: media and maintenance

For DNA transformations and plasmid DNA amplification, two types of chemically competent cells were used. The *Escherichia coli* SURE supercompetent cells (Stratagene), that have reduced recombination potential, were used for transformation and amplification of plasmids containing AAV ITRs and grown at 30℃. The *Escherichia coli* TOP10 competent cells (Life Technologies) were used for all other plasmids and cultured at 37℃. Competent cells were stored at -80˚. Autoclaved LB (Fisher Scientific, 20g/L ddH_20) was used for liquid cultures. Autoclaved LB agar (37g/L ddH_20) set in Sterilin 10cm Petri dishes was used for solid phase growth cultures. Antibiotics were added to autoclaved broth or autoclaved agar cooled to 50°C. Ampicillin (Sigma, 100μg/ml in ddH_2O) or Kanamycin sulfate (Fisher Scientific, 50μg/ml in ddH_2O) were used for selection. Stock preparations of transformed bacteria were kept as glycerol stocks (bacterial pellet resuspended in LB + 10-15% glycerol) at -80℃.

2.1.10 Transformation of competent bacteria

5μl of the DNA insert-vector ligation reaction or 10-50ng of plasmid DNA was used for transformation into competent cells following the manufacturer's protocol. For SURE competent cells, 50μl of competent bacteria were initially incubated with 1μl of β-mercaptoethanol for 10 minutes on ice. DNA was then added and incubated for 30 minutes on ice, followed by a heat shock at 42°C for 45 seconds and further 2 minutes on ice. For Top10 competent cells, 50μl of cells were mixed with DNA and incubated for 30 minutes on ice, followed by a 30 second heat shock at 42℃ and a final 2 minute incubation on ice. In both cases, reactions were then incubated with 250μl of LB at 30℃ for 1 hour and plated on LB agar plates containing the appropriate antibiotic. Plates were incubated upside down at 30℃ or 37°C overnight.

2.1.11 Plasmid DNA amplification and purification – mini preps

Single colonies from transformed or re-streaked bacteria were picked and inoculated in 3ml of LB with antibiotic at the appropriate temperature

overnight in an incubator shaker. The following day, 2ml of the culture were transferred to an eppendorf tube and cells were pelleted in a bench-top centrifuge at 8000 rpm for 4 minutes. DNA was extracted using the QIAGEN mini prep kit. Briefly, cells were re-suspended in QIAGEN P1 buffer and then lysed in 100µl of P2 lysis buffer and mixed by inverting the tube. After a 5-minute incubation at room temperature, chilled P3 neutralisation buffer was added to neutralise the mixture and samples were incubated on ice for 5 minutes. DNA was separated from bacterial debris by centrifugation at 13,000 rpm for 10 minutes in a bench-top centrifuge. Supernatant containing DNA was then transferred to a QIAprep spin column to bind DNA followed by washing with ethanol-containing PE buffer and elution in 40-50µl of ddH_2O. DNA was kept at 4℃ for short term storage or at -20℃ for longer storage.

2.1.12 Plasmid DNA amplification and purification – midi preps

1ml of transformed bacteria culture or 10µl from a glycerol stock were inoculated overnight in 100ml (300-500ml for maxi preps) of LB with antibiotic selection. Cells were harvested by centrifugation at 6,000 rpm for 10 minutes at 4°C and DNA extraction was performed using the NucleoBond® Xtra Midi/Maxi kit (Macherey-Nagel), following the manufacturer's protocol. The general principle behind this plasmid purification protocol is based on a modified alkaline lysis procedure, followed by plasmid DNA binding to a NucleoBond resin under appropriate low salt and pH conditions. RNA, proteins, dyes and low molecular– weight impurities are removed by several wash steps, and the plasmid DNA is finally eluted in a high-salt buffer, concentrated and desalted by isopropanol precipitation, and washed in ethanol. Finally, the clean DNA pellet was allowed to air dry before re-suspending in 100-1,000µl ddH_2O depending on the amount of purified DNA and the desired final concentration. DNA concentration was subsequently measured (see below). Preparations obtained by this method were typically at a concentration of 1µg/µl and stored short-term at 4℃ or long-term at -20℃.

2.1.13 Determination of DNA concentration and DNA sequencing

DNA concentration was determined using a Nanodrop ND-100 Spectrophotometer (Labtech International). Following a blank measurement, 1µl of undiluted DNA preparation was loaded onto the measuring pedestal. The Nanodrop calculates the DNA concentration measuring the sample absorbance at 260nm (OD_{260}) and assuming that 1 OD_{260} unit corresponds to 50µg/ml of dsDNA. The purity of the DNA sample can be assessed by the OD_{260}/OD_{280} ratio, which at ≈1.8 is considered 'pure' for DNA (i.e. free from protein and RNA contamination).

DNA sequencing was performed at Eurofins MWG Operon from 1.5µg of plasmid DNA in 15µl ddH_2O pre-mixed with 2µl of the appropriate primer at 10pmol/µl. Sequencing was performed using the Value Read service in tube format, results were returned on-line and subsequently analysed using the APE DNA analysis software and the NCBI Blast tool.

2.1.14 Plasmids

Here, I present the basic plasmids that were used for this thesis and those that have required cloning. When the cloning strategy is identical for several plasmids, the general strategy is described. For a list of all the plasmids used, please see Table 2 in Appendix 1, pg. 244.

2.1.14.1 Rep-expressing constructs

pCMV-Rep40 (pND229), pCMV-Rep52 (pND230), pCMV-Rep68 Y156F M225G (pND226), and pCMV-Rep78 Y156F M225G (pND227) have been previously described (Dutheil et al., 2014). FLAG- and T7-tagged Rep proteins were generated by cloning of PCR amplified Rep fragments into either the pEGFP-C1 vector (Clontech) containing N-terminal FLAG tag or T7 tag, respectively. Rep K372A mutants were generated by site-directed mutagenesis using primers SS100 and SS101. ZNF truncation mutants were generated by PCR amplification of aa 1-529 (Δ91/87), 1-558 (Δ63), or 1-577 (Δ44), using either pND230 or pND227 as a template. The amplified fragments were then cloned into the N-terminal T7 vector described above using *Hinc*II and *Xma*I sites.

2.1.14.2 BirA*- Rep fusion constructs

Fusion of BirA* to the N-terminus of the Rep proteins was achieved by performing overlapping PCR between PCR-amplified BirA* and each of the PCR-amplified Rep proteins, with the inclusion of a 10x glycine linker. BirA* and each of the Rep sequences were amplified by PCR, after which overlapping PCR was used to fuse the BirA* fragment to the N-terminus of each Rep fragment. The resulting amplicons were cloned into pcDNA3.1+ using *Apa*I sites.

2.1.14.3 WT AAV plasmids

The pDG plasmid is a tool for the production of recombinant and wt AAV. It contains all the adenovirus helper factors necessary for AAV production, as well as the AAV2 Rep and Cap ORFs (Grimm et al., 1998).

The mini-pDG plasmid was derived by Dr Els Henckaerts from pDG and contains only the AAV2 Rep and Cap ORFS. Mutagenesis to generate Rep-K372A for cloning into pAV2 was performed on this plasmid.

The pAV2 plasmid is infectious and contains the full WT AAV genome inserted between two *Bgl*II linkers (Laughlin et al., 1983). pAV2-RepK340H was derived from this plasmid (cloned by Dr. Martino Bardelli in the lab), and pAV2-RepK372A was cloned into this plasmid using the *Sfi*I and *Hind*III fragment generated from the mini-pDG RepK372A variant.

2.1.14.4 RUVBL1, RDBP, TCERG1, CPSF6, KAP1, PP1α, and NIPP1 plasmids

pC1-FLAG-wtKAP1 was a kind gift from Dr. Helen Rowe (UCL) and was used for the cloning of PCR-amplified wtKAP1 into untagged pcDNA3.1+ and the N-terminal T7-tagged vector described above. pEXN-CPSF6 was a kind gift from Dr. Gregory Towers (UCL). FLAG-tagged RUVBL1, RDBP, TCERG1, and PP1α were generated by cloning of PCR-amplified fragments from EST clones obtained from Genome Cube (clones IRAUp969F0818D {RUVBL1}; IRAUp969C0381D {RDBP}; IRCMp5012D00731D {TCERG1}; IRAUp969F0817D {PP1α}) into the N-terminal FLAG-vector described above.

GFP-NIPP1 was obtained from Dr. Angus Lamond (Trinkle-Mulcahy et al., 2001), through Addgene (#44221).

2.2 CELL CULTURE, TRANSFECTIONS, AND INFECTIONS

2.2.1 Cell lines and viruses

Adherent HEK293T and HeLa cells were obtained from the America Tissue Culture Collection (ATCC). HeLa cells are derived from a human epithelial cervical adenocarcinoma (from the patient Henrieta Lacks), while HEK293T cells are human embryonic kidney cells immortalized with adenovirus E1 and E2 genes, and modified to contain the SV40 large T antigen. Cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (GIBCO 41966, Invitrogen) supplemented with 10% fetal bovine serum (FBS) (Invitrogen) plus 100 U/mL Pen/Strep (Gibco). Cells were cultured at 37^{o}C and 5% CO_2 in 10cm dishes (Corning). Cells were passaged every 2-3 days using 1ml TrypLE Express (GIBCO 12605, Invitrogen) so as to maintain the cells at ≤70% confluency.

AAV2 and human adenovirus type 5 (Ad5) were produced and purified as previously described (Zeltner et al., 2010).

2.2.2 Viral titration

Infectious unit (IU) titer for wtAAV2 was determined by infectious center assay, based on similar principles as the standard plaque forming assay. Briefly, HeLa cells were infected with serial dilutions of wtAAV2 stock and coinfected with Ad5 at a multiplicity of infection (MOI) of 25 plaque-forming units (PFU) per cell. 40h after infection, cells were harvested and applied to 0.45µm nylon membranes using a vacuum manifold, after which the membranes were denatured, UV cross-linked, and hybridized to a *rep*-specific probe. The membranes were then washed at high stringency and exposed to a phosphoimager screen for 14-16h. For the titer calculation, each positive signal, corresponding to one infected cell and therefore to one IU, was counted, and the titer was calculated according to the volume and dilution of virus used for the infection. Vector genome (VG)-containing particle titer for wtAAV2 and rAAV was determined by quantitative real-time

PCR, using SYBR Green Jump Start Taq ready mix without $MgCl_2$ (Sigma). Linearized pAV2 (for wtAAV2) or pTRUF11 (for rAAV) was used to set up a standard curve in 10-fold dilutions, ranging from 2ng to 0.002ng, and primers in the *cap* or *GFP* gene were used for wtAAV2 or rAAV, respectively. Plaque forming unit (PFU) titer for Ad5 was determined by standard plaque assay by Marcia Meseck at the Vector Core at Mount Sinai School of Medicine.

2.2.3 Freezing and thawing of cells

For freezing of cells lines, at least 10ml of confluent cells were prepared in 10cm dishes for each desired cryovial. Cells were detached using 1ml of trypsin and pelleted by centrifugation at 1500 rpm for 5 minutes at 4^{o}C, after which the medium was removed and the pellets placed on ice. Cells were then resuspended in 1ml ice-cold freezing buffer containing 50% FBS, 40% culture medium, and 10% DMSO (Sigma) and then immediately transferred to a pre-chilled 1.8ml labeled cryovial. Cryovials were initially stored at -80^{o}C and after 24-48 hours the cryovials were moved to a liquid nitrogen tank. For thawing of cryopreserved cells, vials were thawed at 37^{o}C, transferred to a 10ml tube, and 9ml of pre-warmed culture medium was gently added dropwise to wash the cells. Cells were pelleted at 1200 rpm for 5 minutes and resuspended in 10ml pre-warmed DMEM with 10% FBS. Cells were left to recover in a 6-10cm dish for 24-48 hours.

2.2.4 Transient plasmid transfections

For plasmid DNA transfection, 293T cells were transfected using polyethylenimine (PEI, Polysciences) at a density of 8 x 10^5 cells/ml of medium (e.g. 2ml medium/well of 6-well plate = 1.6 x 10^6 cells). For one well of a 6-well plate, 8µl of PEI was diluted in 40µl serum-free DMEM and allowed to incubate for 5 minutes at room temperature. In a separate tube, 1µg of DNA was diluted in 40µl serum-free DMEM and then combined with the PEI/DMEM mixture and incubated for 20 minutes at room temperature before drop-wise addition to cells. Medium was replaced with fresh DMEM + 10% FBS 6 h after transfection. For larger- and smaller-scale transfections, quantities were scaled accordingly.

2.2.5 siRNA transfections

In a 24-well format, 2 x 10^5 cells were transfected with 50nM (siKAP1.2 and siKAP1.4) or 60-100nM (siRDBP, siRUVBL1, siTCERG1, siCHD3, and siSETDB1) siRNA using 2µl DharmaFECT (Dharmacon) in 50µl Opti-MEM (Gibco). 6h later, medium was replaced with fresh DMEM + 10% FBS. 24 h after transfection, cells were re-plated into 12-well format. 36 h after transfection, cells were subjected to a second transfection as described above, using 4µl DharmaFECT in 100µl Opti-MEM and then used either in a p5 repression assay or AAV2 replication assay. For a list of siRNA sequences used in this thesis, please see Table 3, Appendix 1, pg. 248.

2.2.6 Lentivector production

The HIV-based lentiviral vector pC-SIREN, expressing either a hairpin targeting the 3'UTR of KAP1 (shKAP1; GATCCGCCTGGCTCTG TTCTCTGTCCTTTCAAGAGAAGGACAGAGAACAGAGCCAGGTTTTTTAC GCGTG) or the corresponding empty vector (shEMPTY) were kind gifts from Dr. Helen Rowe (UCL). To produce virus-like particles, 2 x 10^7 293T cells were seeded in a T225 flask and transfected 24h later with 8µg of either pC-SIREN-shKAP1 or pC-SIREN-shEMPTY, 6µg HIV Gag-Pol expression vector, and 6µg VSV-G expression vector using 200µl PEI in a total volume of 2ml serum-free DMEM. Transfection medium was replaced with fresh DMEM + 10% FBS 6h after transfection. Supernatants were harvested 48 and 72 h after transfection, pooled and filtered through a 0.45µM filter, and frozen in 1mL aliquots at -80^{o}C until use.

2.3.7 Lentivector transductions for KAP1 depletion

1 x 10^6 293T cells were transduced in a 6-well format using 1.2-1.6 mL of either shKAP1 or shEMPTY diluted with enough DMEM + 10% FBS to achieve the normal well volume of 2mL. 24h later, the transduction medium was removed, and cells were re-plated at a density of 5 x 10^5 cells/mL for infection with AAV2/Ad5 the next day. For larger- and smaller-scale transductions, quantities were scaled accordingly.

2.2.8 AAV infection and transduction

The same basic principle applied to all infections (using wtAAV2) and transductions (using rAAV). Briefly, cells were infected/transduced at 60-70% confluency in ~2/5 the normal well volume and incubated for 2h at 37^{o}C to ensure maximal adsorption and infection/transduction efficiency. For rAAV transductions, the medium was replenished to the normal culture volume 2h after transduction and cells were harvested 48h after transduction for analysis. For wtAAV2 infections, Ad5 was added to infection medium 2h after AAV2 infection. 1h after Ad5 infection, the medium was replaced with the normal well volume of fresh DMEM + 10% FBS. Cells were harvested ~42h after infection, or when cells displayed optimal cytopathic effect (CPE), at which point all samples were harvested simultaneously. Optimal CPE is defined by cells that display a rounded and enlarged phenotype, as opposed to the normal “star-shaped” morphology of HEK293T cells, and which are beginning to detach but still appear bright and healthy. Cells that had completely lifted by the time of harvest were considered too advanced in the infection cycle and were excluded from analysis.

2.3 DNA, RNA, AND PROTEIN EXTRACTION

2.3.1 Total DNA extraction

For total/genomic DNA extraction, the QIAGEN DNeasy Blood and Tissue Kit was used as per manufacturer’s instructions with one additional step. Briefly, cells were pelleted and washed in PBS, followed by lysis and proteinase K digestion at 56^{o}C to remove all proteins. An additional 10-minute incubation at 70^{o}C was added to ensure complete lysis of stable AAV2 particles. The DNA isolation/purification was performed using a silica-based DNA purification in spin-columns using the kit buffers; DNA was eluted in 200µl ddH_2O.

2.3.2 Total RNA extraction

RNA extraction was performed using the RNeasy Mini Kit (QIAGEN), following the manufacturer's instruction, with some minor modifications. Cells were pelleted and washed in PBS, followed by lysis in 350µl of the supplied buffer. Lysates were homogenized using a syringe and 21G needle and column purified. RNA was eluted in 50µl RNAse-free water and aliquots were taken to determine the concentration using the Nanodrop spectrophotometer. For DNaseI treatment, 10µg of RNA were diluted in 85µl RNAse-free water and mixed with 10µl RDD buffer (supplied in kit) and 5µl DNaseI stock solution (QIAGEN DNAseI kit; prepared as per manufacturer's protocol) and allowed to incubate for 20 minutes at room temperature. The DNaseI-treated RNA was then re-purified using the RNeasy kit following the RNA clean-up protocol. RNA was then stored at -80^{o}C.

2.3.3 Protein extraction

Volumes indicated are for one well of a 6 well plate. Cells were harvested and pelleted at 1500 rpm for 5 minutes, washed in PBS, and pelleted again. Cells were then lysed in ~50µl RIPA buffer (25mM Tris HCl pH 7.4, 150mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, 1X protease inhibitor cocktail [Complete, Roche]) on ice for 10 minutes. Lysates were clarified by centrifugation at 13,000 rpm for 10 minutes at 4^{o}C and then transferred to a clean 1.5ml tube. Proteins were either used right away or were stored at -80^{o}C until use.

2.4 DETECTION OF DNA AND RNA BY qPCR

2.4.1 Quantitative PCR (qPCR) – absolute quantification

For analysis of viral replication, total DNA was extracted using the Qiagen DNAeasy Blood and Tissue DNA extraction kit as described in section 3.3.1. Viral DNA was quantified by qPCR using the SYBR Green JumpStart Taq ReadyMix for qPCR (Sigma-Aldrich) according to the manufacturer's protocol using an ABI PRISM system (Applied Biosystems). The $MgCl_2$ and primer concentrations were optimized previously to 4mM and 0.25µM respectively (Zeltner et al., 2010). Linearized pDG was used to

prepare a standard curve for the quantification of AAV2 and Ad5, ranging from 1 x 10^2 – 1 x 10^8 molecules of dsDNA in 10-fold dilutions. Similarly, 293T genomic DNA was used to prepare a standard curve of 0.2 – 5ng of DNA in 5-fold dilutions for the quantification of cyclophilin. Total extracted DNA was diluted 100-fold, and 1µl was used for quantification using primers specific for AAV2 *cap*, Ad5 *100kd protein*, and *cyclophilin.* The cycling parameters were the following: The cycling parameters were the following: A 2 minute initial denaturation step at 94°C, followed by 40 cycles of a denaturation step at 94°C for 15 seconds, an annealing step at 58°C for 20 seconds, and an elongation step at 72°C for 1 minute.

2.4.2 Reverse transcription

For RNA analysis, RNA was reverse-transcribed using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) following the manufacturer's protocol. 1-2µg of RNA were diluted to a final volume of 10µl using RNAse-free water and then added to 10µl of a 2X RT master mix prepared by mixing the following provided reagents; 2µl 10X RT buffer, 0.8µl dNTP mix (100mM), 2µl 10X random primers, 1µl MultiScribe™ Reverse Transcriptase, and 4.2µl nuclease-free ddH_2O. RT negative controls were prepared in parallel by excluding the 1µl MultiScribe™ Reverse Transcriptase and using 5.2µl nuclease-free ddH_2O. Reverse transcription was performed using a thermal cycler with the following conditions: 10 minutes at 25°C, 2 hours at 37°C, and 5 minutes at 85°C. cDNA was stored at 4°C for short-term storage or at -20°C for long-term storage, and was used for relative quantification by qPCR (see below).

2.4.3 qPCR - relative quantification of cDNA

3'FAM-5'TAMRA-conjugated qPCR probes and qPCR primers designed to bind *rep* or *cap* were obtained from MWG Eurofins Operon (see table for details). cDNA was diluted 10-fold, and qPCR was performed using TaqMan universal PCR master mix (Life Technologies) and the custom primer-probe mix in a total reaction volume of 10µl. Primers were used at a final concentration of 900nM, and probes at 250mM. Relative expression levels were determined with the ΔΔCT quantification method (Schmittgen

and Livak, 2008), using 18S rRNA (Taqman Pre-developed assay reagents, human 18S rRNA, Applied Biosystems) as a housekeeping reference gene.

2.5 DETECTION OF KAP1 BINDING TO THE AAV2 GENOME

2.5.1 Chromatin immunoprecipitation (ChIP)

Cells were cross-linked in their medium in 1% formaldehyde (10 minutes at room temperature) and quenched with 0.125M glycine (5 minutes at room temperature) before being lysed in lysis buffer (1ml per $1x10^8$ cells; 50mM Tris-HCl, pH 8, 10mM EDTA, 1% SDS, 1x protease inhibitors) for 10 minutes on ice. Lysates were sonicated on a Branson Sonifier 250 to obtain 200- to 500-bp fragments (15 x 30-second cycles with 90-second intervals, output ~2). 10µl of the lysate was used to assess sonication efficiency by reverse cross-linking for 15' at 95°C and then incubating with RNAse A for 30 minutes at 37°C. DNA was extracted and visualized on a 1.5% agarose gel. The remaining lysate was clarified at 13,000 rpm for 10 minutes at 4 degrees. The equivalent of $2x10^6$ cells was diluted 25-fold in RIPA buffer (50mM Tris pH8, 150mM NaCl, 2mM EDTA, pH 8, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 1x protease inhibitors) and pre-cleared with 80µl protein G agarose beads (pre-blocked in 0.1mg/mL BSA for 30 seconds) for 2h on a rotator at 4°C. For the immunoprecipitation, primary antibodies were added to the lysates and incubated for 1h on a rotator at 4°C (5µg IgG [Abcam; ab37415], 4µg H3K9me3 [Abcam; ab8893], 10µg KAP1 [Abcam; ab10483]), before adding 80µl pre-blocked beads and incubating overnight on a rotator at 4°C. Beads were harvested and washed 4 times in RIPA buffer, 4 times in high salt wash (20mM Tris-HCl, pH 8, 1mM EDTA, 500mM NaCl, 0.5% NP-40, 1X protease inhibitors), 4 times in TE buffer (10mM Tris-HCl, pH8, 1mM EDTA), and eluted in 160ml elution buffer (100mM $HaHCO_3$, 1% SDS) for 15 minutes at 30°C. Cross-links were reversed by adding NaCl to a final concentration of 0.2M and incubating overnight at 67°C. Eluates were then incubated with 2µl RNase A (10mg/mL) and 2µl proteinase K (20mg/mL) at 45°C for 1h. DNA was purified using a PCR purification kit (Qiagen) and analyzed by qPCR using primers specific for *GAPDH*, *ZNF180*, *ZNF274*, or various regions of the AAV2 genome.

2.5.2 ChIP-qPCR

Purified chromatin was diluted 10-fold and quantified by real-time PCR using the SYBR Green JumpStart Taq ReadyMix for qPCR (Sigma-Aldrich) using an ABI PRISM system (Applied Biosystems). Primer sequences for *GAPDH*, *ZNF180*, *ZNF274*, and the various regions of the AAV2 genome are listed in Table 1, Appendix 1, pg. 239. CT values for "10% input" were adjusted by subtracting 3.322 cycles to correct for the 10-fold dilution factor (https://www.thermofisher.com/us/en/home/life-science/epigenetics-noncoding-rna-research/chromatin-remodeling/chromatin-immunoprecipitation-chip/chip-analysis.html). Percent input was then calculated as follows: 100 x 2^-(CT of adjusted 10% input - CT of ChIP-ed DNA). Percent input for each antibody was then normalized to values for IgG to calculate final fold enrichment.

2.6 PROTEIN DETECTION

2.6.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

2-4µl of protein extracts in RIPA buffer (in section 2.3.3) were mixed with 2µl of 6X Laemmli loading buffer (0.8g SDS, 5mL Tris pH 6.8, 5mL glycerol, 5% beta-mercaptoethanol, trace of bromophenol blue) and adjusted to a total volume of 12µl with ddH_2O. Samples were denatured by boiling at 95^oC for 10 minutes and then separated on 6% or 12% polyacrylamide mini-gels. The percentage was determined by altering the volume of 40% Acrylamide (Acrylamide:Bis-Acrylamide 29:1, Fisher Scientific) and ddH_2O added to a solution of 1.5M Tris-HCl pH8.8, 0.4% SDS, 3.3µl/mL of 10% ammonium persulphate (APS, Sigma), and 0.6µl/ml of N,N,N'N' – Tertramethylethylenediamine (TEMED, Sigma). Stacking gels consisted of 4.5% Acrylamide, 0.5M Tris-HCl pH6.8, 0.4% SDS, 5µl/ml 10% APS, 1µl/mL TEMED, and ddH_2O. Samples were loaded into gels immersed in running buffer (1X Tris-glycine and 1% SDS in ddH_2O prepared from a 10X Tris-glycine solution [288g glycine and 61g Tris base dissolved in 2L ddH_2O]) and run at 80V until the samples reached the separating gel, after which the

voltage was increased to 120V until samples reached the bottom of the gel. 5µl of Precision Plus Dual Color Standard (Bio-Rad) was run in parallel and used as a marker for molecular weights.

2.6.2 Western blotting and protein detection

Proteins resolved in SDS-PAGE were transferred to a nitrocellulose 0.45µM membrane (Hybond-C Extra nitrocellulose, Amersham Biosciences) using a BioRad Mini Trans-Blot transfer system with transfer buffer (100mL 10X Tris-Glycine, 200mL methanol, 700mL ddH_2O) at 16V overnight or at 100V on ice for 1-2 hours. Membranes were blocked with either 5% (w/v) skimmed dried milk, or 2.5% (w/v) bovine serum albumine (BSA) for phospho-antibodies, in PBST (0.1% Tween-20, Sigma, in 1X PBS) for 45 minutes at RT. Primary antibodies were added directly to the blocking medium and allowed to incubate for 2h at RT or overnight at 4^{o}C, after which membranes were washed 3 x 10 minutes in PBST and incubated with HRP-conjugated anti-mouse or anti-rabbit IgG (BioRad) for 1h at RT. After a further 3 x 10-minute washes in PBST, membranes were developed using the West Pico Enhanced Chemiluminescent Substrate (Thermo Scientific) prepared according to the manufacturer's protocol. After a 5-minute incubation in a total volume of 2mL of substrate, membranes were placed between two clear plastic sheets and visualized using an Image Quant LAS4000 (GE Healthcare) chemiluminescence imaging system. For a list of antibodies used in this thesis, please see Table 4, Appendix 1, pg. 249.

2.6.3 Fluorescence-assisted cell sorting (FACS) analysis

GFP transfected or transduced cells were trypsinised, washed in PBS, passed through a cell strainer to ensure single-cell separation, and analysed for protein expression on a BD FACSCanto II flow cytometer using the BD FACSAria II program (BD Biosciences). Results were analysed using FlowJo software.

2.7 ANALYSIS OF PROTEIN-PROTEIN INTERACTIONS

2.7.1 Proximity-dependent biotin identification (BioID)

Ten 10cm dishes (293T) per condition were transfected using 6µg DNA and 50µl PEI in 500µl SF medium per dish. 6h post-transfection, the medium was replaced with fresh DMEM + 10% FBS, and D-Biotin (Life Technologies) was added to a final concentration of 100µM. 48h post-transfection, cells were washed twice in cold PBS to remove remaining biotin and lysed in 1 ml lysis buffer (50 mM Tris, pH 7.4, 500 mM NaCl, 0.4% SDS, 5 mM EDTA, 1 mM DTT, and 1x Complete protease inhibitor). Lysates were sonicated for 20 seconds, on a Branson Sonifier 250, output ~2, followed by the addition of Triton-X-100 to a final concentration of 2%. Lysates were sonicated another 20 seconds after which 1ml of 50mM Tris HCL pH 7.4 was added before a third sonication of 20 seconds. Lysates were clarified at 13,000 rpm for 5 minutes at 4^{o}C, and the supernatants were incubated with 200µl avidin-agarose beads (Fisher) for 3h at 4^{o}C. The beads were then collected and washed two times by gravity filtration in 1ml wash buffer 1 (2% SDS). This was repeated once with wash buffer 2 (0.1% deoxycholate, 1% Triton X-100, 500mM NaCl, 1 mM EDTA, and 50mM Hepes, pH 7.5), once with wash buffer 3 (250mM NaCl, 0.5% NP-40, 0.5% deoxycholate, 1mM EDTA, and 10mM Tris, pH 8.1), and twice with wash buffer 4 (50mM Tris, pH 7.4, and 50mM NaCl). Proteins were eluted from beads by boiling in 50µl 2X Laemmli SDS buffer saturated with biotin at 98^{o}C for 10 minutes and sent for analysis by LC-MS/MS.

2.7.2 Liquid chromatography mass spectrometry (LC-MS/MS)

Mass spectrometry was performed by Steven Lynham at the KCL proteomics facility at Denmark Hill. In-gel reduction, alkylation and digestion with trypsin were performed prior to subsequent analysis by mass spectrometry. Cysteine residues were reduced with dithiothreitol and derivatised by treatment with iodoacetamide to form stable carbamidomethyl derivatives. Trypsin digestion was carried out overnight at RT after initial incubation at 37^{o}C for 2h. Peptides were extracted from the gel pieces by a series of acetonitrile and aqueous washes. The extract was pooled with the

initial supernatant and lyophilised. Each sample was then resuspended in 10μL of 50mM ammonium bicarbonate and analysed by LC-MS/MS. Chromatographic separations were performed using an EASY NanoLC system (ThermoFisherScientific, UK). Peptides were resolved by reversed phase chromatography on a 75μm C18 column using a three step linear gradient of acetonitrile in 0.1% formic acid. The gradient was delivered to elute the peptides at a flow rate of 300 nL/min over 120 min. The eluate was ionised by electrospray ionisation using an Orbitrap Velos Pro (ThermoFisherScientific, UK) operating under Xcalibur v2.2. The instrument was programmed to acquire in automated data-dependent switching mode, selecting precursor ions based on their intensity for sequencing by collision-induced fragmentation using a Top20 CID method. The MS/MS analyses were conducted using collision energy profiles that were chosen based on the mass-to-charge ratio (*m/z*) and the charge state of the peptide. Raw mass spectrometry data were processed into peak list files using Proteome Discoverer (ThermoScientific; v1.4.0.288)

2.7.2 Immunoprecipitation (IP)

293T cells were transfected in a 6-well format with 200ng of Rep-expressing constructs, and 250ng of FLAG-PP1α or FLAG-GFP using 8μl PEI in 80μl serum-free DMEM. 2 wells per condition were used for IP. 48h after transfection, cells were trypsinised in 200μl TrypleE and harvested using ice cold 1X PBS + 10% FBS. Cells were pelleted at 1500 rpm for 5 minutes at 4^{o}C and then lysed in RIPA buffer for 10 minutes on ice. Lysates were clarified at 1000 x *g* for 10 minutes at 4^{o}C and transferred to clean 1.5ml tubes. 50μl of lysate was separated at this point to use as an input control, and the remaining lysate was incubated with the appropriate antibody for 1.5h on a rotator at 4^{o}C. 40μl of protein G agarose beads (Pierce) were added and incubated for a further 3h. Beads were washed 4 times in 500μl RIPA buffer, and proteins were eluted from beads by boiling at 95^{o}C for 10 minutes in 60μl 2X Laemmli buffer. 50μl of 2X Laemmli buffer was added to the input control samples and boiled as described for the beads.

2.7.3 Cross-linked immunoprecipitation (CL-IP)

In a 6-well format, 293T cells were transiently transfected with the appropriate constructs using 8µl PEI in 80ul SF medium (200ng Rep-expressing constructs; 750ng FLAG-GFP, FLAG-KAP1, GFP-NIPP1, FLAG-RDBP, FLAG-RUVBL1, FLAG-TCERG1, or CPSF6). 48 hours after transfection, cells were harvested in trypsin, washed in cold PBS + 10% FBS, and fixed in 350µl 0.05% formaldehyde for 10 minutes at 37^{o}C. Cells were washed again in cold PBS and incubated in 350µl 0.125M glycine, pH 7, for 5 minutes at RT in order to quench the reaction, before being lysed in 500µl cross-linked IP buffer (150mM NaCl, 10mM HEPES pH 7, 6mM $MgCl_2$, 2mM DTT, 10% glycerol, 1X protease inhibitors, 200µM sodium orthovanadate) on ice for 10 minutes. Lysates were subjected to three 10-second cycles of sonication on a Branson Sonifier 250, output ~2, and clarified for 10 minutes at 4^{o}C. 20µl of lysate was kept for an input control, and 420µl was used for IP. For the IP, 40µl protein G agarose beads (Pierce) per sample were washed 3 times in cross-linked IP buffer and incubated with 2µg mouse anti-FLAG (Sigma, F7425) or anti-T7 antibody, or 7µg anti-CPSF6 antibody for 1.5h on a rotator at 4^{o}C before being added to 420µl of cell lysate and incubated a further 3-4 h. Beads were harvested and washed 4 times in 500 µl RIPA buffer, and cross-links were reversed in 25µl reverse cross-link buffer (10mM EDTA, 5mM DTT, 1% SDS) at 65^{o}C for 45 minutes. Proteins were eluted from beads by adding 30µl of 2X Laemmli loading buffer and boiling at 95^{o}C for 10 minutes before being analyzed by western blot.

2.8 IMAGING

2.8.1 Transfection and infection

293T cells were seeded at a density of 1x10^{5} on poly-L-lysine (Sigma) coated coverslips (autoclaved coverslips were placed in 24-well plate, incubated for 10 minutes with 400µl poly-L-lysine, and washed once with 1X PBS) in 24-well plates the day prior to transfection. 4h prior to transfection, DMSO (Sigma) or ATMi (Stratech Scientific Ltd) was added to the appropriate wells to a final concentration of 10µM. Cells were then

transfected with 20ng of empty vector or pRep78-GFP (a kind gift from Prof. Cornel Fraefel at the University of Zurich) using 2µl Lipofectamine 2000 in 50µl serum-free Opti-MEM. The next day, cells were infected with Ad5 (2 PFU/cell) in a total volume of 160µl for 1h, after which the medium was replaced for fresh DMEM + 10% FBS. Cells were fixed 24h after Ad5 infection for staining and visualization (see below).

2.8.2 Staining and visualization

Cells were fixed in 4% PFA for 10 minutes at RT, washed 3 times for 5 minutes in PBS, permeabilized in 0.1% Triton-X-100 for 10 minutes at RT, and washed again 3 times for 5 minutes in PBS. Cells were then incubated with primary antibody (anti-p-KAP1-S824; 1:1000) diluted in PBS + 1% BSA for 2h at RT, washed, and then incubated with secondary antibody (Biolegend; rabbit IgG2b-AlexaFluor 594, 1 µg/ml (1:1000)) diluted in PBS for 1h at RT. Cells were then washed a final time and mounted in Prolong Gold Antifade Reagent (Invitrogen). Images were visualized using an Eclipse Ti-E Inverted confocal microscope and analyzed using NIS Elements C software.

2.9 ANALYSIS OF AAV2 LIFE CYCLE

2.9.1 Analysis of WT AAV2 infection in KAP1-depleted cells

Control or KAP1-depleted 293T cells (section 3.2.3) were infected at ~80% confluency by adding AAV2 at the stated MOI in ~2/5 the normal well volume. 2h after AAV2 infection, Ad5 was added at an MOI of 2 PFU/cell, and medium was replaced with fresh DMEM + 10% FBS 1h after Ad5 infection. Cells were harvested for DNA, RNA, and protein extraction ~42h after infection, or when they displayed optimal cytopathic effect (CPE) (described in section 2.2.7).

2.9.2 Anlysis of AAV2-Rep-K372A infection

8 x 10^6 293T cells (untreated, control, or KAP1-depleted where appropriate) seeded in a 10cm dish were transfected with 16µg of pAV2-Rep-WT, pAV2-Rep-K340H, or pAV2-Rep-K372A using 90µl of PEI and 500µl SF medium. 4h later, cells were infected with 2 PFU/cell of Ad5, and

the medium was replaced with fresh DMEM + 10% FBS 2h after Ad5 infection. Cells were harvested for DNA, RNA, and protein extraction 72h after transfection.

2.10 ANALYSIS OF PHOSPHORYLATED KAP1-S824

Phosphorylation of KAP1-S824 was investigated in 293T cells that were either infected with AAV2 and/or Ad5, or transfected with various Rep-expressing constructs using linear PEI. Where relevant, cells were pretreated with either DMSO or 10μM ATMi 4h prior to infection/transfection, and inhibitors were maintained throughout. Cells were harvested for western blot analysis 27h after infection/transfection, unless otherwise stated.

Chapter 3. Application of BioID for the investigation of interactions between the small AAV2 Rep proteins and components of the general cellular transcriptional machinery.

3.1 Introduction

The duality of the biphasic AAV life cycle requires an exceptional degree of complexity from what is genomically speaking a simple virus. As such, the simplicity of the AAV genome belies a complex strategy for transcriptional regulation involving the interplay of host cellular, helper virus, and AAV Rep proteins. In the absence of helper virus, the three AAV promoters are repressed through the combined efforts of the cellular factors YY1 and MLTF, and the large AAV Rep proteins (Beaton et al., 1989; Chang et al., 1989; Kyöstiö et al., 1994; Shi et al., 1991). AAV recruits YY1 and MLTF to binding sites in its p5 promoter where they mediate repression in the absence of helper virus factors; this consequently represses all three promoters as p19 and p40 are transactivated through p5 activity. Simultaneous binding of the large Rep proteins to both a repressive RBS in the p5 promoter and an activating RBS in the ITR (Labow et al., 1986) acts to balance repression such that the basal levels of *rep* expression necessary to mediate latency-associated transcriptional repression and viral integration are still maintained. Upon coinfection, helper virus factors such as Ad E1A interact with YY1 and MLTF, relieving repression of p5 (Chang et al., 1989; Shi et al., 1991; Tratschin et al., 1986) and leading to the transactivation of all three AAV promoters (Redemann et al., 1989) in a manner that supports higher levels of expression from p40 than from either p5 or p19 (Mouw and Pintel, 2000). As infection progresses, the disproportionate accumulation of spliced Rep variants results in an autoregulatory feedback loop that serves to maintain optimal levels of gene expression (Pereira et al., 1997).

Transcriptional repression of the AAV promoters by Rep in the absence of helper virus has been shown to occur via two independent mechanisms: (1) one mechanism present only in the large Rep proteins, which relies on the presence of a functional RBS within the target site promoter, and (2) repression by all four Rep proteins, which requires only an

intact NTP-binding site within the Rep proteins (Kyostio et al., 1995; Kyöstiö et al., 1994). Structural and biochemical studies suggest that binding of the RBS in the p5 promoter by the OBD-containing large Rep proteins followed by the assembly of Rep helicase complexes may structurally hinder either the start of transcription or the binding of necessary transcriptional activators (Chiorini et al., 1994; Dutheil et al., 2014; Zarate-Perez et al., 2012). Optimal repression of p5 is also dependent on a functional Rep NTP-binding motif, suggesting that the efficiency of helicase formation and/or its DNA affinity may also play a role (Dutheil et al., 2014). In contrast, repression of the p19 promoter is dependent only on the NTP-binding site. Furthermore, the small Rep proteins, which do not contain an OBD and which only form transient helicase complexes, are also able to mediate transcriptional repression of all three AAV promoters, indicating the potential reliance of this pathway on interactions between cellular factors and the Rep ATPase domain (Dutheil et al., 2014; Kyöstiö et al., 1994).

Intriguingly, Rep-mediated repression has also been observed for various heterologous promoters. All four of the Rep proteins were shown to repress transcription from the *PPP1R12C* promoter within the target cellular integration site, *AAVS1,* via mechanisms similar to what has been observed for the AAV p5 promoter (Dutheil et al., 2014). Similarly to p5, repression of the *PPP1R12C* promoter was more efficient by the large Rep proteins than by the small Rep proteins, which could be attributed to binding of the cellular RBS by the N-terminal OBD of the large Rep proteins. However, Rep40 and Rep52 were also shown to repress both p5 and the *PPP1R12C* promoter in a manner dependent on the NTP-binding motif. Furthermore, the *PPP1R12C* promoter displays bidirectional repression activity, and repression of the antisense promoter by all four Rep proteins was dependent only on the NTP-binding motif (Dutheil et al., 2014).

Similarly, Rep-mediated repression has been observed for various heterologous viral promoters. Rep78, Rep68, and Rep52 are able to repress transcription from the human immunodeficiency virus long terminal repeat (HIV LTR), the human papillomavirus type 18 upstream regulatory region (HPV18 URR), and the Ad major late transcription promoter (AdMLP) (Hörer et al., 1995; Needham et al., 2006). Repression of the HIV LTR and the

HPV18 URR was dependent on the NTP-binding motif, and mutational analysis of the HPV18 URR revealed that several *cis* acting sequences were necessary for this repression, supporting the notion that Rep protein interactions with cellular transcription factors are involved (Hörer et al., 1995). Direct binding of Rep78 to the TAR element of the HIV LTR was also demonstrated however, suggesting that repression of HIV may occur via both mechanisms (Batchu and Hermonat, 1995). Repression of AdMLP was shown to be dependent on both the NTP-binding site as well as the direct interaction of the large Reps with a 55 bp sequence in the AdMLP and TBP bound adjacent to this site, provoking the hypothesis that Rep interferes with PIC formation of RNA pol II transcription (Needham et al., 2006). In addition, Rep78 was shown to down-regulate expression of several proto-oncogene promoters, such as those of c-Myc, c-Fos, H-Ras, and c-sis/platelet derived growth factor B (Hermonat, 1994, 1991; Wonderling and Owens, 1996).

While a role for the Rep ATPase domain in the repression of multiple promoters has been clearly demonstrated, the exact mechanism through which this occurs has yet to be fully described. One possible candidate involved in repression by the Rep proteins is the human positive cofactor 4, or PC4. PC4 was first identified as a transcriptional coactivator by virtue of its ability to activate transcription in the presence of various sequence-specific TF (Ge and Roeder, 1994; Kretzschmar et al., 1994), however it was later demonstrated to also have an intrinsic repressive quality arising from its capacity to structurally stabilize the formation of RNA pol II PIC, thereby preventing the activation of elongation (Malik et al., 1998). This repression is subsequently reversed by phosphorylation/inactivation of PC4 in the presence of the general transcription factors TFIID and TFIIH. Both Rep68 and Rep78 were shown to bind PC4 in a manner dependent on an intact NTP-binding site *in vivo* (Weger et al., 1999). The accumulation of active, nonphosphorylated PC4, induced through overexpression, resulted in downregulation of all three AAV2 promoters suggesting that PC4 is involved in the repression of AAV2 gene expression (Weger et al., 1999). Given the role for active, nonphosphorylated PC4 in PIC stabilization and the fact that AAV promoters contain TATA boxes, it is conceivable that repressive PC4-mediated PIC at AAV promoters are stabilized through interactions with Rep.

Preliminary knockdown experiments in our lab however showed no effect of PC4 depletion on AAV transcription.

Similarly, repression could be the result of interactions between Rep and other general components of the RNA pol II holoenzyme, including the general transcription factors TFIIA, -B, -D, -E, -F, and -H. For example, Rep interaction with TBP, a component of TFIID, has been shown to inhibit adenovirus major late transcription factor (Hernonat et al., 1998; Needham et al., 2006). Other possibilities include – (1) inhibition of gene expression through interference with splicing mechanisms, and (2) modulation of DNA structure and/or higher order chromatin structure. Several of the known Rep interaction partners have roles in modifying DNA structure, including PC4, HMG1, and SP1, and a comprehensive analysis of proteins in Rep complexes by Nash et al. identified numerous splicing /RNA-binding proteins. Additionally, this work identified several proteins involved in regulating transcription through the modulation of higher order chromatin structure, including TIF1α and RuvB-like 1. The possibility that Rep may interfere with such essential transcriptional processes as the formation of RNA pol II PICs poses the question of how AAV can remain innocuous during latency. Even more intriguingly, the observation that the Rep proteins can regulate expression from the *PPP1R12C* promoter suggests that AAV may have coevolved with its host in such a way as to maintain control of its own integration site (Dutheil et al., 2014). In either case, knowledge of which cellular proteins are involved in mediating repression by the Rep proteins will have broad implications for our understanding of AAV interactions with both host and helper virus.

Common methods for the screening of novel protein interactions can be limited in their scope by an inherent dependence on both the persistence and strength of protein interactions. Here, we investigated interactions between the small Rep proteins and cellular proteins in order to identify a potential mechanism for Rep-mediated repression of cellular promoters using a screening method known as Biotin Identification (BioID) (Roux, 2012). BioID exploits the fusion of a promiscuous biotin ligase, BirA*, to a bait protein in order to trigger proximity-dependent biotinylation of neighboring proteins, thus allowing for the identification of a much broader scope of

protein associations than achievable with conventional affinity purification. This method was developed and validated in a study in which BirA* was fused to the well-characterized intermediate filament protein, lamin-A (LaA) (Roux et al., 2012). This study demonstrated that the BirA*-LaA fusion protein was correctly targeted to the nuclear lamina, where it led to the specific biotinylation of proteins associated with the nuclear envelope (NE), in contrast to the nucleoplasmic biotinylation that occurs in the presence of BirA* alone. The investigators were able to identify multiple known interaction partners of LaA, as well as an uncharacterized interaction partner later named SLAP75 (Roux et al., 2012). One key attribute of this approach lies in its detection of protein-protein interactions within their natural cellular environment, thus avoiding problems associated with the incorrect folding or post translational modification of bait/prey proteins, as can be the case in the commonly used yeast two hybrid (Y2H) approach. In addition, BioID sidesteps the issue of protein solubilization that is associated with traditional affinity purification methods, which may not be compatible with preserving weak protein interactions. As the biotinylation of vicinal proteins occurs prior to solubilization, BioID allows for the detection of weak, transient, and indirect interactions.

3.2 Results

3.2.1 BirA*-Rep fusion proteins are stable and retain transcriptional regulation and biotin ligase activity.

To generate BirA*-Rep bait proteins for BioID, we fused cMyc-tagged BirA* (BirA*) to the N-terminus of both Rep40 and Rep52 by overlapping PCR. The equivalent fusions were generated between BirA* and the repression-deficient NTP-binding mutants, Rep40-K340H and Rep52-K340H, to use as negative controls for transcriptional repression (Chejanovsky and Carter, 1990). By comparing results obtained from the screens using BirA*-Rep40 and BirA*-Rep52 with their NTP-binding mutant counterparts, we hoped to narrow down our list of potential candidates. To ensure each fusion protein was stable and expressed at similar levels, 293T cells were transiently transfected with plasmids expressing fusion or WT Rep proteins. Cells were harvested 48 h after transfection and analyzed by western blotting using anti-Rep and anti-c-Myc antibodies for the detection of BirA* (Figure 12A). Rep40 and BirA*-Rep40 expression was equivalent, as was that of Rep40-K340H and BirA*-Rep40-K340H. The higher expression levels displayed by the K340H mutants was expected due to the decreased capacity of these proteins to autoregulate. Both BirA*-Rep52 and BirA*-Rep52-K340H expressed less efficiently than their unfused counterparts, however the lack of any observable degradation products suggested that the proteins were stable. Blotting with anti-c-Myc confirmed BirA* stability. Additionally, c-Myc expression levels were equivalent for all four of the fusion constructs.

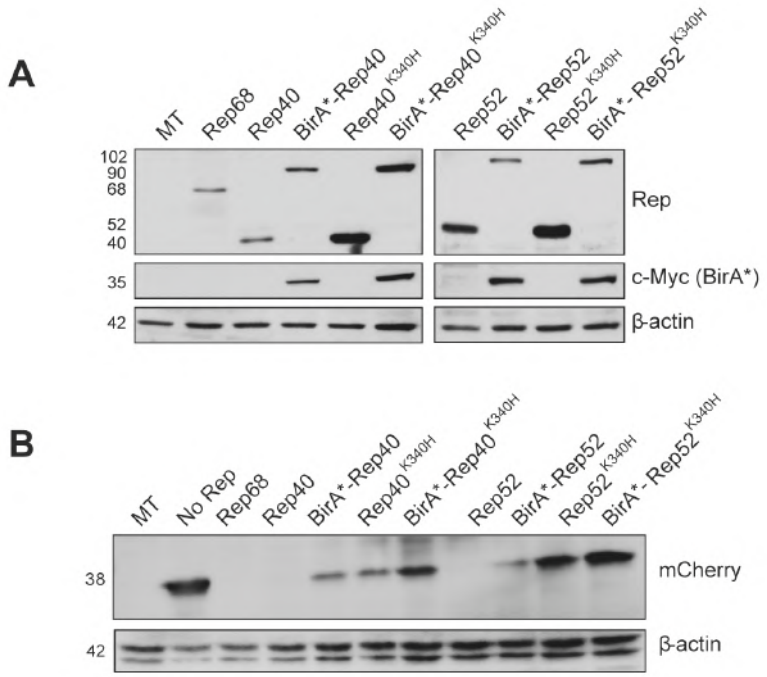


Figure 12. Functional validation of Rep activity in BirA*-Rep fusion proteins
(A) 293T cells were either mock transfected (MT) or transfected with the indicated Rep or BirA*-Rep fusion protein, and protein expression levels were analyzed by western blotting 48h after transfection using anti-Rep and anti-cMyc antibodies. **(B)** The regulatory activity of the fusion proteins was determined by their ability to repress the AAV2 p5 promoter. 293T cells transfected with a p5-mCherry reporter construct and each of the BirA*-Rep fusion constructs, or each WT Rep proteins, were harvested 48h after transfection, and p5 activity was determined by western blotting for mCherry. Rep68 was used as an additional control for repression.

To confirm that the fusion proteins retained their regulatory function, we assessed their ability to repress mCherry expression from a reporter construct in which mCherry is driven by the AAV2 p5 promoter. p5-mCherry was transiently cotransfected in 293T cells with each of the fusion proteins, or the respective Rep proteins for comparison. Rep68 was used as an additional positive control for repression. Cells were harvested 48 h after transfection, and p5 activity was assessed by western blot for mCherry (Figure 12B). Repression of p5 by BirA*-Rep40 was diminished as compared to Rep40 and appeared similar in efficiency to that of the repression-deficient Rep40-K340H mutant. However, BirA*-Rep40 repressed p5 more efficiently than BirA*-Rep40-K340H, suggesting that the overall activity of the Rep40 fusion proteins was diminished rather than that the fusion proteins were not

functional. While repression of p5 by BirA*-Rep52 was also marginally less efficient than by Rep52, the effect was much less apparent than with the Rep40-based constructs.

Having confirmed that the Rep proteins retained their regulatory function, we next sought to confirm the activity of BirA*. Each construct was transiently transfected into 293T cells, with or without the addition of 100μM free biotin, and cells were harvested 48h after transfection for protein extraction. Biotinylated proteins were enriched using streptavidin agarose beads and were visualized by western blotting using avidin-peroxidase (Figure 13). A clear abundance of biotinylation was evident with all four fusion proteins, confirming that the biotin ligase activity of BirA* remained intact. Surprisingly however, no clear differences were observable between the K340H mutant and wild-type Rep fusion proteins. The strength and abundance of an interaction however is not necessarily reflective of its biological significance, and it is possible that differences in bands not visible by western blot would be detected by mass spectrometry analysis.

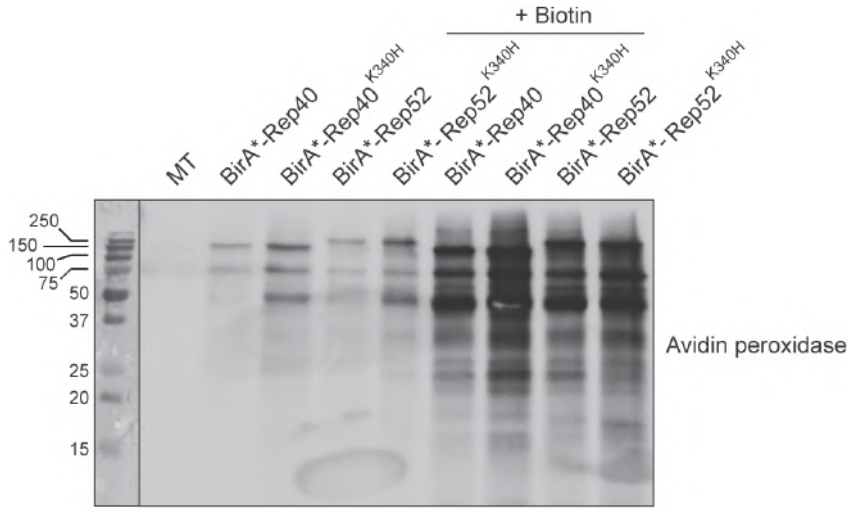


Figure 13. BirA* activity in BirA*-Rep fusion proteins
293T cells transfected with each of the BirA*-Rep fusion proteins, in the presence or absence of free biotin, were harvested 48h after transfection. Biotinylated proteins were enriched using avidin agarose beads and biotin ligase activity of BirA* was determined by western blotting using avidin peroxidase.

3.2.2 BioID identifies various candidates of interest as novel potential interaction partners for the AAV2 Rep proteins.

We next generated samples of biotinylated proteins for analysis by LC-MS/MS. 293T cells were transiently transfected with either BirA*-Rep52 or BirA*-Rep52-K340H as described above and lysed 48h after transfection. Mock-transfected (MT) 293T cells were processed in parallel as a control. For these experiments, ten confluent 10-cm dishes were used per condition. Biotinylated proteins were captured using streptavidin agarose beads, rigorously washed, and bound proteins were analyzed by mass spectrometry. The efficiency of biotinylation was verified by western blotting as described above prior to LC-MS/MS analysis (Figure 14). A list of all the proteins identified by the two pull downs is shown in Table 5 in Appendix 2, pg. 250. As expected, virtually no proteins were identified in the MT sample, with the exception of trypsin and serum albumin. After applying a peptide identity confidence threshold of 95% and a minimum requirement for two identified peptides to the list of hits, proteins were categorized based on their cellular function (Figure 15).

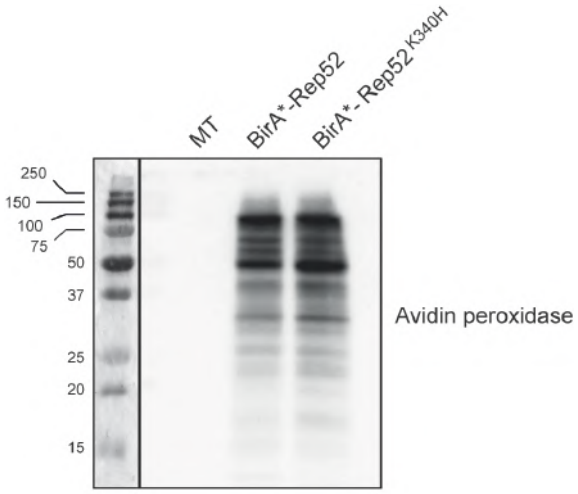


Figure 14. Qualitative analysis of biotinylated samples for mass spectrometry analysis.
Biotinylation in samples for LC-MS/MS analysis was verified by western blotting using avidin peroxidase.

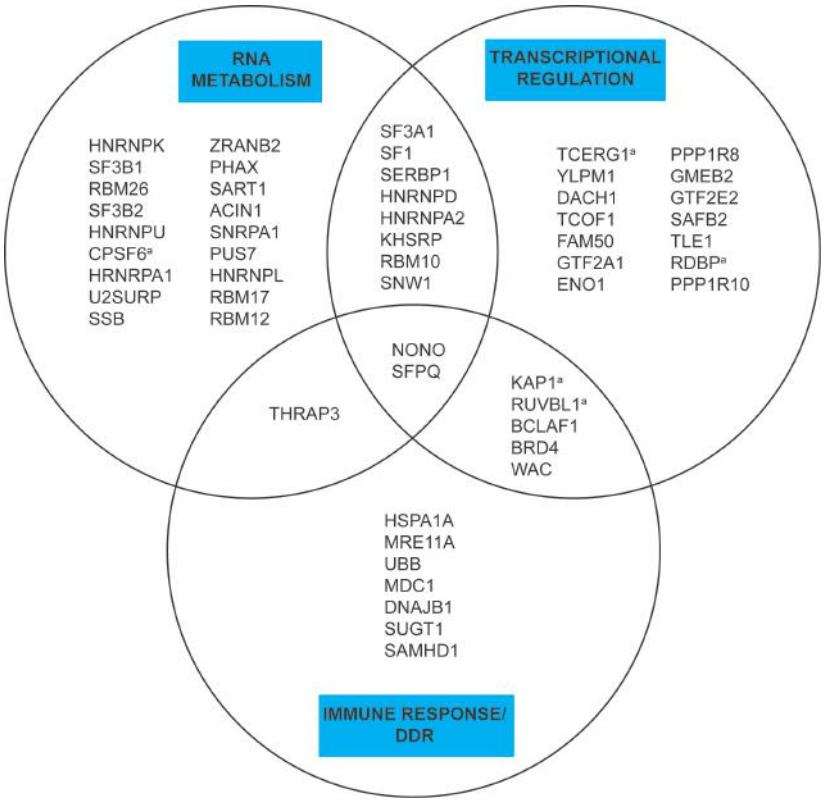


Figure 15. Proteins identified in BirA*-Rep52 and BirA*-Rep52^{K340H} BioID screens.
Proteins were categorized based on cellular function after applying a peptide identity threshold of 95% and a minimum requirement of 2 peptides.
[a] Denotes candidates chosen for further downstream analysis.

A multitude of factors involved in RNA metabolism and transcriptional regulation were identified, as well as several proteins involved in the DNA damage response. Surprisingly however, there was no clear difference in interaction profiles between BirA*-Rep52 and BirA*-Rep52-K340H; it appeared in fact that BirA*-Rep52-K340H supported more interactions than its wild-type counterpart, rendering it unusable as the control it was intended to be. As a result, we considered proteins identified with either BirA*-Rep52 or BirA*-Rep52-K340H for downstream analysis. Several of our hits were in fact proteins previously identified as Rep interaction partners by other groups, such as RUVBL1, RDBP, SNW1, and MDC1 (Cervelli et al., 2008; Nash et al., 2009), lending support to the quality and coherence of our results. In the absence of our intended negative control however, we selected candidates on which to focus based on their cellular function in transcriptional regulation as well as on any previously known association with viral elements. RUVBL1 is thought to act as a key assembly factor for a

variety of biological processes, including transcriptional regulation, and is a known component of the ION80 complex involved in chromatin remodeling (Chen et al., 2011). Furthermore, it has been shown to interact with YY1 (Wu et al., 2007), a transcription factor known to regulate AAV2 p5. The negative elongation factor RDBP, one of five subunits of the NELF complex, stabilizes proximal-promoter pausing during transcription by RNA Pol II (Yamaguchi et al., 1999) and thus represents another candidate potentially involved in Rep-mediated transcriptional regulation. TCERG1 and SF1 cooperate to physically couple the process of transcriptional elongation with RNA splicing; TCERG1 binds the CTD of RNA Pol II and recruits SF1, which then targets nascent transcripts for degradation (Goldstrohm et al., 2001). Both TCERG1 and SF1 were identified in our screen. Of additional interest, TCERG1 has previously been linked to Tat-dependent transcriptional regulation of HIV-1 (Coiras et al., 2013). The pre-mRNA processing factor CPSF6 also caught our interest, however not for its role in RNA metabolism but rather for its recently identified role in HIV-1 infection. During the early stages of HIV-1 infection, CPSF6 binds the HIV-1 capsid, preventing premature release of viral RNA and onset of reverse transcription, thus avoiding immune detection by cytoplasmic pattern recognition receptors (PRRs).

Of most interest however was the identification of the transcriptional corepressor KAP1, which will be discussed further in chapters 4, 5, and 6. These top hits were later confirmed in screens performed with BirA*-Rep40/Rep40-K340H, BirA*-Rep68-K340H, and BirA*-Rep78-K340H. Complete lists of peptides identified by BioID in these screens are shown in Tables 6 and 7 in Appendix 2, pg. 255 and 258.

3.2.3 Interactions between RUVBL1, RDBP, and TCERG1, and Rep52 are not essential for Rep-mediated transcriptional regulation.

To confirm the interaction between Rep52 and the chosen candidates of interest, we first repeated BioID in the context of exogenously expressed FLAG-tagged candidate proteins. cDNA of RUVBL1, RDBP, and TCERG1 derived from EST clones was cloned into expression vectors containing an N-terminal FLAG tag, and expression of each construct was verified by transient transfection of 293T cells followed by western blot analysis of cell

lysates with anti-FLAG antibody. FLAG-tagged RDBP (expected size ~50kD) and RUVBL1 (expected size ~50kD) expressed efficiently and without any observable degradation of the tagged proteins (Figure 16, A and B). FLAG-TCERG1 (expected size ~150kD) expression was less consistent, with all three clones tested showing considerable degradation (Figure 16C). Furthermore, two prominent bands were visible, one at ~150kD and one at ~200kD. Given that a faint band at 150kD was also visible in the MT sample, the larger than expected 200kD band likely represents TCERG1. Certain downstream analyses were therefore not possible for TCERG1, including the repeat of BioID and cross-linked immunoprecipiation (CL-IP) experiments. For future work, it will be necessary to experiment with alternative tags as well as compare the amino- versus carboxy-terminal location of each tag in order to achieve efficient expression of TCERG1.

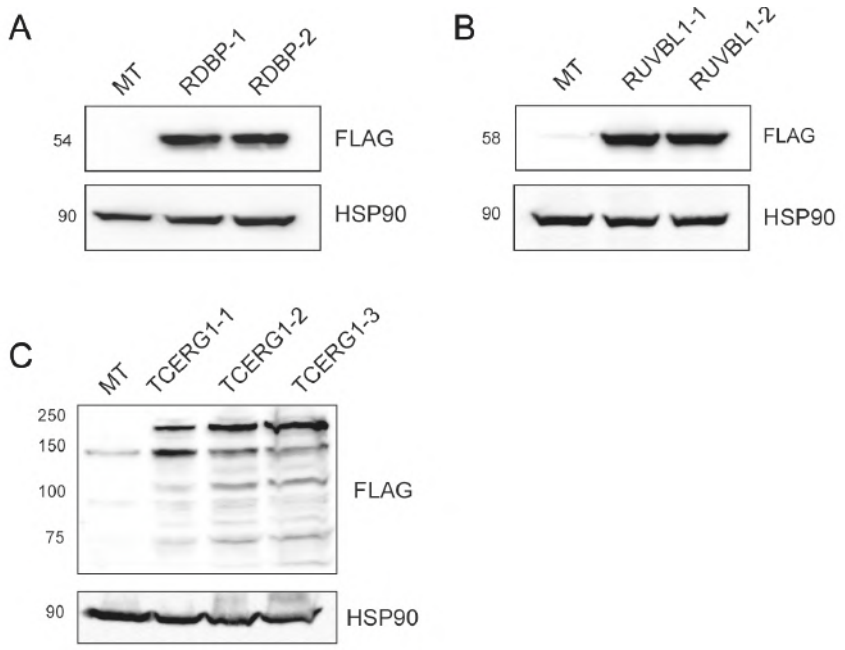


Figure 16. Expression of FLAG-tagged RDBP, RUVBL1, and TCERG1.
293T cells transfected with multiple clones of FLAG-tagged RDBP **(A)**, RUVBL1 **(B)**, or TCERG1 **(C)** were harvested 48h after transfection. Expression of FLAG-tagged proteins was analyzed by western blotting.

To repeat the BioID, FLAG-RUVBL1 and FLAG-RDBP were expressed with either BirA*-Rep52 or BirA*-Rep52-K340H in 293T cells in the presence of free biotin. Biotinylated proteins were purified 48 h after transfection and analyzed by western bloting using anti-FLAG and anti-Rep antibodies. FLAG-RUVBL1 and FLAG-RDBP were each readily detected among the purified proteins in the context of BirA*-Rep52 and BirA*-Rep52-K340H, confirming the putative interaction between these candidates and Rep52 (Figure 17, A and B). We next used co-IP experiments to gain more insight into the nature of these interactions. Conventional co-IPs were unsuccessful, suggesting that the interactions identified by BioID were indirect, weak, and/or transient. By cross-linking cells prior to harvest however (CL-IP), we were able to confirm physical associations between Rep52 and both RDBP and RUVBL1. FLAG-RDBP, FLAG-RUVBL1, or a FLAG-Y14 control was coexpressed with Rep52 in 293T cells. 48 h after transfection, cells were cross-linked in 0.05% formaldehyde, and FLAG-tagged proteins were immunoprecipitated from the lysates. Rep52 was readily detectable in both the FLAG-RDBP and FLAG-RUVBL1 pull-down fractions. (Figure 17, C and D). Although weak bands were visible at ~50kD in the FLAG-Y14 fraction, these appeared to stem from the light chain of the rabbit FLAG antibody used for IP still present in the lysates, as the use of mouse FLAG antibody for western blotting greatly diminished these bands without affecting the strength of bands in either FLAG-RDBP or FLAG-RUVBL1 fractions (data not shown).

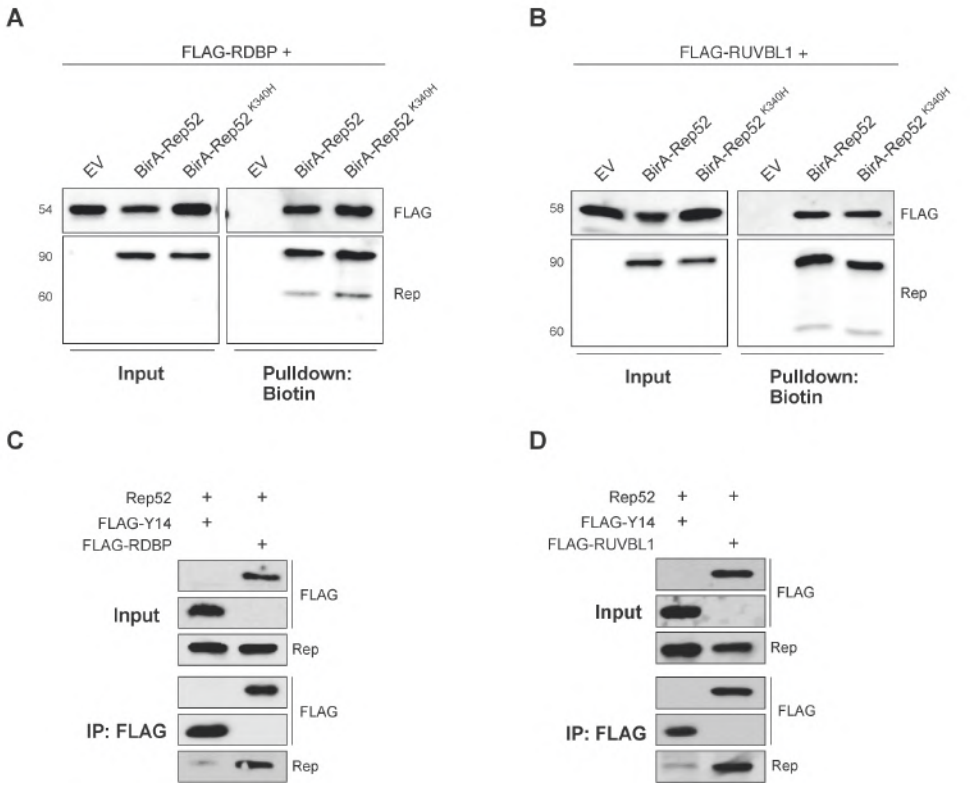


Figure 17. Validation of the physical interaction of Rep52 with RDBP and RUVBL1. **(A-B)** Purified biotinylated proteins from 293T cells transfected with empty vector (EV), BirA*-Rep52, or BirA*-Rep52^{K340H} and FLAG-RDBP **(A)** or FLAG-RUVBL1 **(B)** were analyzed for Rep and FLAG by western blotting. **(C, D)** Cross-linked IP for FLAG-tagged proteins from lysates of 293T cells expressing Rep52 with a FLAG-Y14 control, FLAG-RDBP **(C)** or FLAG-RUVBL1 **(D)**.

To determine the potential contribution of RUVBL1 and RDBP to Rep-mediated transcriptional regulation, we next assessed the ability of Rep52 to regulate expression from the AAV2 p5 promoter in the context of RUVBL1 and RDBP depletion. Although we were unable to validate the physical association between Rep52 and TCERG1, we also included TCERG1 in this set of experiments to functionally validate this candidate. 293T cells were transfected with siRNAs targeting RUVBL1, RDBP, TCERG1, or luciferase as a control. Knockdown cells were then cotransfected with a p5-mCherry reporter construct, described in section 4.2.1, and increasing concentrations of Rep52. Activity of the p5 promoter was determined 48 h later by western blotting for mCherry. Repression of p5 by Rep52 appeared diminished in RUVBL1 knockdown cells in the presence of the lowest concentration of

Rep52, however repression was equally efficient at higher concentrations of Rep52, suggesting that RUVBL1 is not involved in Rep-mediated transcriptional repression (Figure 18A). TCERG1 depletion also had no effect on repression of p5 by Rep52, however reduced expression of mCherry in TCERG1 knockdown cells transfected with p5-mCherry alone suggested a possible role for this candidate in the positive regulation of the AAV2 p5 promoter (Figure 18B). Despite extensive optimization efforts, we were unable to achieve efficient knockdown of RDBP (Figure 18C), and it is therefore impossible to conclude at this time what the contribution of RDBP might be to Rep-mediated transcriptional repression.

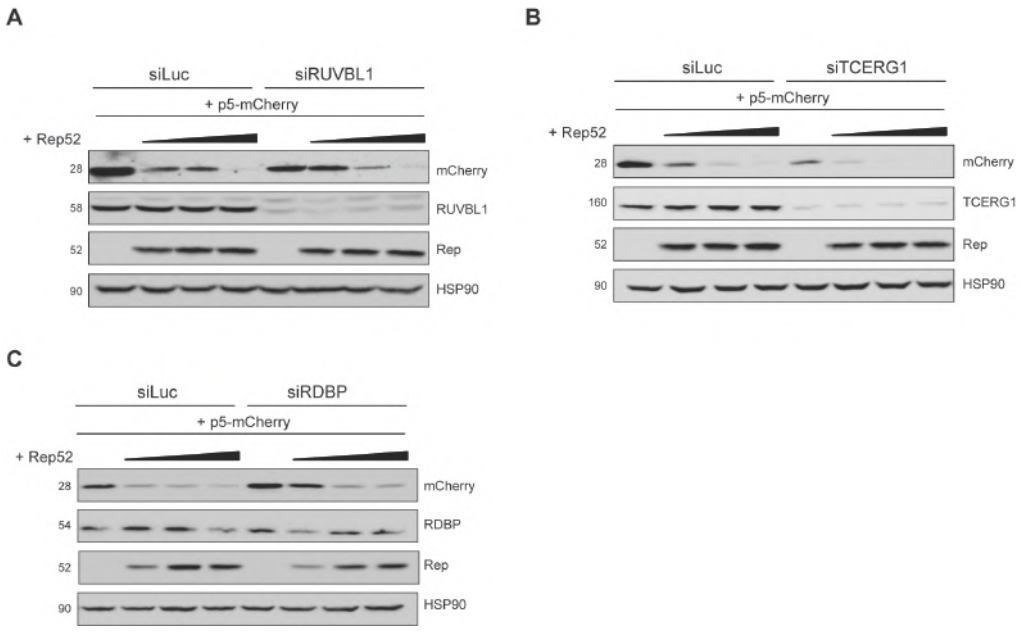


Figure 18. Functional analysis of RUVBL1, RDBP, and TCERG1.
To determine the role of each candidate in Rep-mediated transcriptional regulation, we assessed the ability of Rep52 to repress the AAV2 p5 promoter in the absence of each candidate. 293T cells treated with siRNAs targeting RUVBL1 **(A)**, TCERG1 **(B)**, RDBP **(C)**, or luciferase as a control were transfected with a reporter construct in which mCherry is driven by the AAV2 p5 promoter, and increasing concentrations of Rep52. Cells were harvested 48h after transfection, and p5 activity was determined by western blotting for mCherry.

3.2.4 Potential role for interactions between CPSF6 and capsid-associated Rep in AAV2 replication.

During HIV-1 infection, CPSF6 has been shown to bind the incoming viral capsid, preventing the premature release of viral RNA and onset of reverse transcription, thus helping to avoid immune sensing by cytoplasmic pattern recognition receptors. There is also evidence to suggest that this interaction helps to guide the viral particle to the nuclear pore. We questioned whether a similar mechanism could be occurring during AAV2 infection, and if this might account for observed differences in nuclear trafficking between the various AAV serotypes. We first repeated BioID in the context of exogenously expressed CPSF6 to validate the putative interaction between Rep52 and CPSF6, as described above for RDBP and RUVBL1. This confirmed the presence of CPSF6 in purified biotinylated proteins generated by both BirA*-Rep52 and BirA*-Rep52-K340H (Figure 19A). Conventional co-IP experiments failed to reveal a physical interaction between Rep52 and CPSF6, therefore we performed CL-IP experiments as described for RDBP and RUVBL1. CPSF6 was coexpressed with GFP, Rep52, or Rep52-K340H, and was immunoprecipitated from cell lysates 48h after transfection (Figure 19B). Interestingly, CPSF6 appeared to physically interact exclusively with Rep52-K340H, however input expression levels for Rep52-K340H were much greater than for Rep52, which may account for this result.

We then asked whether Rep could be interacting with CPSF6 at the viral capsid. The Rep proteins have been shown to interact with the viral capsid. Rep78 covalently bound to the 5' end of viral genomes is thought to target viral genomes to the capsid for encapsidation by forming a complex with capsid proteins (Dubielzig et al., 1999; Prasad and Trempe, 1995; Prasad et al., 1997). In addition, all four Rep proteins have been shown to interact directly with capsid proteins, forming helicase complexes at the 5-fold pore necessary for genome encapsidation (Bleker et al., 2006; King et al., 2001).

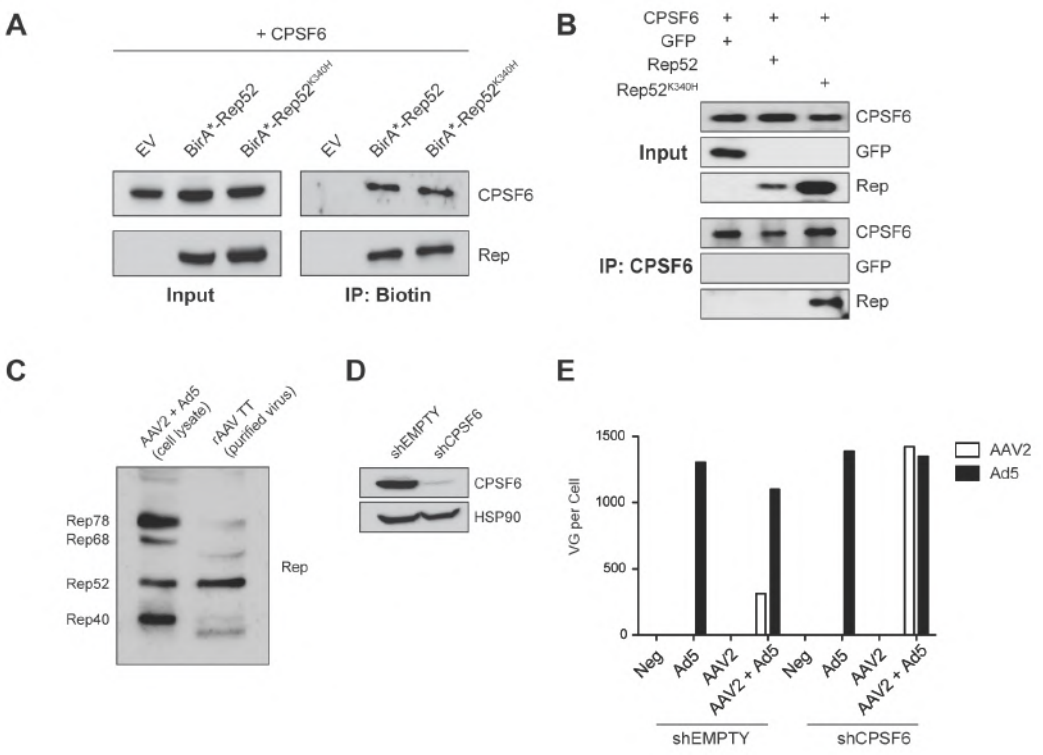


Figure 19. Validation of the Rep-CPSF6 interaction
(A) Biotinylated proteins purified from cells expressing CPSF6 and BirA*-Rep52, BirA*-Rep52^{K340H}, or empty vector (EV) in the presence of free biotin were analyzed by western blotting for the presence of CPSF6 and Rep. **(B)** Co-IP for CPSF6 from lysates of 293T cells expressing CPSF6 with GFP, Rep52, or Rep52^{K340H}. **(C)** 1 x 10^{11} particles of purified rAAV TT were analyzed by western blotting for the presence of Rep. Lysates obtained from 293T cells infected with AAV2 and Ad5 were run in parallel. **(D)** 293T cells transduced with either a retroviral vector expressing a shRNA targeting CPSF6 (shCPSF6), or an empty vector control (shEMPTY) were analyzed by western blotting 48h after transduction for efficiency of CPSF6 depletion. **(E)** Control (shEMPTY) or CPSF6 depleted (shCPSF6) 293T cells were infected with Ad5 at an MOI of 5 PFU/cell, AAV2 at an MOI of 10 IU/cell, or coinfected with Ad5 and AAV2 at the aforementioned MOIs to initiate lytic AAV2 replication. Cells were harvested 48h after infection for real time qPCR analysis of viral genomes.

It is unclear however how long these interactions last, and if Rep proteins are still associated with capsids upon reinfection. To determine if Rep bound to the AAV capsid might play a role during early stages of infection, we first looked to see whether Rep could be detected in association with purified AAV particles. Proteins from 1 x 10^{11} particles of purified recombinant AAV2 (rAAV2) were subjected to SDS-PAGE and probed by western blotting with anti-Rep antibody. Proteins from lysates of 293T cells

coinfected with WT AAV2 and Ad5 were used for comparison. Surprisingly, all four Rep proteins were detectable with purified rAAV2 particles, and Rep52 was highly abundant compared to Rep40, Rep68, and Rep78 (Figure 19C). Also of interest was the apparent shift in band size for Rep40 and Rep68, indicating a potential cleavage or other modification of these proteins.

We then reasoned that if CPSF6 were involved in the trafficking of AAV2 to the nucleus, depletion of CPSF6 would have an observable effect on AAV2 replication efficiency. To explore this possibility, we performed genome replication experiments in cells depleted for CPSF6. 293T cells were transduced with retroviral vectors expressing either an shRNA targeting CPSF6, or the empty vector as a control (Figure 19D). 48h after transduction cells were infected with Ad5 alone, AAV2 alone, or coinfected to initiate productive replication. Cells were harvested for total DNA extraction approximately 42h after infection, or when cells displayed optimal cytopathic effect (CPE), and replication efficiency was determined by quantitative PCR (qPCR). In cells coinfected with Ad5 and AAV2, we observed a 4.5-fold enhancement in AAV2 replication in cells depleted for CPSF6 as compared to controls (Figure 19E). AAV2 replication efficiency in general however was poor. During an efficient lytic cycle, AAV2 replication should progressively overpower that of Ad5. While some repression of Ad5 replication was apparent in coinfected control cells, the effect was minimal and was not detectable in cells depleted for CPSF6. Furthermore, replication of Ad5 in coinfected cells was as efficient as that of AAV2. Future work with respect to CPSF6 replication experiments will require optimization in order to achieve the expected replication efficiency for AAV2 in control cells. This may require testing alternative strategies for CPSF6 depletion, such as siRNA or lentiviral-based shRNA vectors, as well as various multiplicities of infection for both Ad5 and AAV2.

3.3 Discussion

This work was carried out in an effort to further elucidate the mechanism of Rep-mediated repression that has been observed for heterologous cellular and viral promoters (Dutheil et al., 2014; Hermonat, 1994, 1991; Hörer et al., 1995; Needham et al., 2006; Wonderling and Owens, 1996). While repression of the AAV p5 promoter is dependent on direct binding of the large Rep proteins to the p5 RBS as well as a functional NTP-binding site within the Rep ATPase domain, repression of p19, p40, and several heterologous promoters is dependent only on the presence of a functional NTP-binding site (Kyostio et al., 1995; Kyöstiö et al., 1994), indicating the potential reliance of this pathway on interactions between Rep and components of the cellular transcriptional machinery. In an effort to identify potentially relevant cellular proteins, we used the screening method BioID (Roux et al., 2012) to screen for novel interaction partners of the small Rep proteins. The candidate proteins TCERG1, RUVBL1, RDBP, CPSF6, and KAP1 were chosen for downstream analysis based on their respective roles in transcriptional regulation and/or antiviral activity. CL-IP experiments confirmed the physical association of Rep52 with RUVBL1 and RDBP, however no role could be established for either of these proteins, or for TCERG1, in the negative regulation of the AAV p5 promoter by Rep52. A physical interaction between CPSF6 and Rep52-K340H was also confirmed by CL-IP, and depletion of CPSF6 led to an unexpected enhancement in AAV2 replication in AAV2 and Ad5-coinfected cells.

As with any approach, there are certain limitations of BioID that must be kept in mind. Clearly, the success of BioID is dependent upon the fusion protein maintaining its correct targeting and assembly characteristics. In this work, we generated fusions between BirA* and the N-termini of Rep40 and Rep52, as well as the repression deficient NTP-binding mutants, Rep40-K340H and Rep52-K340H. Each fusion protein displayed stable expression, albeit with variable efficiency of expression. As observed for the WT Rep proteins, the NTP-binding mutant fusion proteins expressed more efficiently than their WT fusion protein counterparts. It is important to consider the possible effect of Rep concentration on protein interactions, as Rep expression is normally highly regulated, both during AAV latency and lytic

replication. In addition, the large Rep proteins display concentration-dependent oligomerization properties that are potentially essential for supporting protein interactions (Bardelli et al., 2016; Zarate-Perez et al., 2012, 2013), and which may be adversely affected by non-physiological levels of Rep expression. Although the small Rep proteins are monomeric in solution (Zarate-Perez et al., 2012) and have only been demonstrated to form transient dimers in the presence of ATP and a DNA template (Smith and Kotin, 1998), it is still possible that both (1) the vast difference in expression levels between the BirA*-Rep fusion proteins and physiological levels of Rep found during AAV infection, and (2) the difference in expression between the WT and NTP-binding mutant fusion proteins may have affected the results of our screen. The use of native AAV promoters might have proven beneficial in this respect, however the highly regulated nature of AAV promoters would likely result in levels of expression too low for effective biotinylation.

Using a transfection-based reporter assay in which mCherry is driven by the AAV p5 promoter (Dutheil et al., 2014), we were able to show that the fusion proteins maintained their capacity to repress p5, albeit with lower efficiency, suggesting that their repressive activity was not compromised. While we felt that this method of validation was sufficient to proceed with our screen, it would have been prudent to more specifically assess ATPase and helicase activities, as well as to verify the localization of the BirA*-Rep fusion proteins by immunofluorescence. Indeed, confocal microscopy could have shed light on the observed diminished repressive activity of the fusion proteins by revealing any potential differences in the efficiency of nuclear translocation. However, as BirA* and the Rep proteins in the absence of helper virus coinfection display similar pan-nuclear patterns of expression, it would have been difficult to visually demonstrate changes in BirA*-mediated patterns of biotinylation resulting from the fusion of Rep to BirA*. This last point further provokes the question of what different protein interactions might have been observed either in the presence of Ad5 coinfection, or with the addition of AAV DNA as a template for the nucleation of protein interactions.

A second limitation of BioID arises from the mechanism of biotinylation itself. The efficacy of BioID is contingent upon the presence of primary amines, namely lysines, within the target protein. As a result, neither the strength nor absence of biotinylation can be directly translated to the importance of a particular interaction. For this reason, BioID screens should be used to identify candidates that can be subsequently investigated systematically or in a hypothesis-based manner. Furthermore, the covalent attachment of biotin to primary amines results in a loss of charge at these sites, which can adversely affect protein interactions, function, and/or secondary modifications. This point may explain our observation that, rather than act as the intended negative controls, the K340H fusion proteins in fact supported more interactions than their WT counterparts. It is conceivable that BirA* biotin ligase activity, which is lysine specific (Roux et al., 2012), resulted in the autobiotinylation of K340 in Rep52. This would effectively render Rep52 equivalent in function to Rep52-K340H, while additionally obstructing any potential protein interactions dependent on the structural integrity of this site. On the other hand, Rep52-K340H could be unaffected by this biotin modification and, being a catalytic mutant and not a structural one, would still be able to support normal protein interactions. Therefore, we considered proteins identified with both WT and K340H fusion proteins for downstream analysis. While this left us without a proper negative control, it is worth noting that the proteins identified in our screen varied greatly from those identified both by Roux et al, and from those identified by various groups in our department employing BioID in the context of unrelated viral proteins. Furthermore, several proteins previously identified as Rep interaction partners in other screens (Cervelli et al., 2008; Nash et al., 2009; Pegoraro et al., 2006) were also identified here, lending support to the quality and coherence of our BioID screen. Nevertheless, it would be valuable to repeat these experiments with the inclusion of a BirA* control. While BirA* would lead to high levels of nonspecific biotinylation of nuclear proteins, identifying enrichment over background of biotinylation in the presence of BirA*-Rep would likely enable the identification of true interaction partners.

In the absence of an RBS within the target promoter, repression by the Rep proteins is dependent only on the presence of an intact Rep NTP-

binding site (Kyostio et al., 1995; Kyöstiö et al., 1994), which suggests that the correct conformation of the ATPase domain, and its oligomerization in the context of the large Rep proteins, is necessary, potentially to support interactions between Rep and components of the cellular transcriptional machinery (Dutheil et al., 2014; Hörer et al., 1995; Needham et al., 2006). Rep is known to interact with various cellular transcription factors, such as E2F1, c-Jun, TBP, and PC4, in a manner shown to down-regulate promoter expression (Batchu et al., 2001; Hernonat et al., 1998; Needham et al., 2006; Prasad et al., 2003; Weger et al., 1999). Conversely, Rep-Sp1 interactions are necessary for the transactivation of the AAV p19 and p40 promoters (Hermonat et al., 1996; Pereira and Muzyczka, 1997a, 1997b). Furthermore, interactions between Rep, HMG1, and the TAF1/Set complex suggest that changes in chromatin structure may also be important for Rep-mediated transcriptional regulation (Costello et al., 1997; Pegoraro et al., 2006). Indeed, changes in higher order chromatin structure are increasingly recognized as being essential for both cellular and viral transcriptional regulation (Knipe et al., 2013; Li et al., 2007a).

Our screen identified a multitude of factors involved in RNA metabolism and transcriptional regulation, as well the regulation of chromatin structure, including TCERG1, RUVBL1, RDBP, CPSF6, and KAP1, which were chosen for further downstream analysis based on their roles in transcriptional regulation and/or their previous association with viral elements. Transcription elongation regulator 1 (TCERG1/CA150) is a nuclear protein first identified as a component of the RNA pol II holoenzyme in a screen for cellular proteins involved in Tat-dependent HIV-1 transcriptional regulation (Sune et al., 1997), however it was later suggested to act as a cofactor for HIV-1 transcriptional elongation by increasing the rate of paused RNA pol II transcription through the phosphorylation of Ser2 in the RNA pol II C terminal domain (CTD) (Coiras et al., 2013). In the context of cellular gene regulation, the role of TCERG1 is to couple the processes of transcriptional elongation and mRNA processing by binding the phosphorylated CTD of actively transcribing RNA pol II and recruiting transcription and splicing factors (Carty et al., 2000; Pearson et al., 2008), such as the splicing factor 1

(SF1) (Goldstrohm et al., 2001). SF1 is then thought to target nascent transcripts, leading to the inhibition of elongation (Goldstrohm et al., 2001).

We were unable to confirm the physical interaction between Rep and TCERG1 due to inefficient expression from CMV-FLAG-TCERG1, however repression of p5 by Rep52 in the absence of TCERG1 was unaffected, suggesting that the TCERG1/SF1 mechanism is not involved in Rep-mediated repression. Alternatively, the reporter-based p5 repression assay does not fully recapitulate certain required physiological aspects of AAV infection. DNA structures such as the ITRs may be influential in facilitating protein interactions. Additionally, physiological levels of Rep during AAV infection are highly regulated, and therefore over-expression of Rep in this assay may also impede certain Rep functions. Finally, protein depletion by siRNA is not 100% effective, and it is possible that the minimal levels of protein remaining are sufficient to support Rep-mediated repression. Interestingly however, expression of mCherry was greatly diminished in the absence of TCERG1. Rep expression on the other hand was unaffected, suggesting that TCERG1 may play a role in the specific positive regulation of p5. It is possible that TCERG1 may play a similar role for p5 as has been observed for Tat-activated HIV-1 LTR transcription, where it appears to phosphorylate the CTD of RNA pol II, thereby releasing it from promoter proximally paused positions (Coiras et al., 2013; Sune and Garcia-Blanco, 1999; Sune et al., 1997) . It would be interesting to determine RNA pol II occupancy on Rep-regulated promoters as well as the phosphorylation status of its CTD in the presence and absence of TCERG1. Alternatively, splicing of nascent mRNA by TCERG1/SF1 may be necessary for productive transcription from p5.

RDBP is an RNA-binding component of the negative elongation factor (NELF) protein complex, which has a major role in regulating proximal promoter pausing through interactions with RNA and RNA pol II (Yamaguchi et al., 1999, 2002). Additionally, studies investigating the regulation of HIV-1 transcription revealed that RDBP binding to the TAR element downstream from paused RNA pol II inhibits expression at the level of early elongation (Pagano et al., 2014). While we were able to confirm a physical association between RDBP and Rep52, we cannot conclude anything with respect to its

role in Rep-mediated repression, as we were unable to achieve efficient depletion of RDBP despite extensive optimization efforts. The particular siRNA sequences used in this set of experiments were obtained from the publications of other groups working on these candidates. We failed to consider the fact that efficiency would likely vary between cell types, which could explain our difficulties in depleting RDBP. In the future, commercial siRNA pools will first be tested, after which the particular siRNA sequences mediating efficient knockdown will be experimentally derived and used for further analysis.

RUVBL1, along with the closely related RUVBL2, has been associated with various cellular processes, including transcriptional regulation and the cellular DDR, through its role in facilitating the assembly of multiple chromatin remodeling complexes that utilize the energy of ATP hydrolysis to reorganize chromatin and increase access to transcription factors and DNA repair factors (Jha et al., 2008; Jin et al., 2005; Mizuguchi et al., 2004; Shen et al., 2000; Torreira et al., 2008). In particular, RUVBL1/RUVBL2 complexes nucleate the assembly of the INO80 complex involved in nucleosome remodeling, a component of which is the transcription factor YY1 (Cai et al., 2007; Jin et al., 2005; Wu et al., 2007), also known to bind the AAV p5 promoter and thus providing a potential link between RUVBL1 and p5 activity. A member of the AAA+ family of helicases, RUVBL1 also bears striking similarities to the Rep proteins, which led us to speculate that Rep might interact with and inhibit RUVBL1, thereby interfering with the assembly of crucial activating chromatin remodeling complexes, such as the TIP60 acetyltransferase complex. The fact that YY1 is a component of INO80 also presented the tempting possibility that INO80 might be necessary for YY1-mediated activation of p5 in the presence of helper virus coinfection, and that Rep counteracts this activity during latency. Again however, depletion of RUVBL1 had no effect on Rep-mediated repression of p5. As with TCERG1 however, expression from p5 appeared marginally less efficient with RUVBL1 depletion, suggesting a potential role for these chromatin-remodeling complexes in the positive regulation of p5.

CPSF6 is a component of the cleavage factor Im complex (CFIm) that plays a key role in pre-mRNA 3'-processing, including 3'-end poly(A) site

cleavage and poly(A) addition (Brown and Gilmartin, 2003; Kim et al., 2010; Rueegsegger et al., 1998). CPSF6 is thought to bind to RNA substrates and promote RNA looping (Yang et al., 2011), although the exact role for CPSF6 in the CFIm complex is not clear. Interestingly however, CPSF6 has been shown to be necessary for infection of monocyte-derived macrophages (MDM) by HIV-1 (Rasaiyaah et al., 2013). It is suggested that CPSF6 binding to HIV-1 CA effectively cloaks reverse transcription, allowing for the evasion of innate immune sensors in primary human macrophages. We asked if CPSF6 might also bind AAV capsids and play a role in nuclear localization.

We could only confirm an interaction between CPSF6 and Rep52-K340H, which might be explained by the much greater abundance of Rep52-K340H than Rep52 in cell lysates; furthermore, if the interaction between Rep and CPSF6 is capsid-dependent, then detecting a physical interaction in the absence of capsid proteins would prove difficult. We were able to show that Rep proteins were indeed still associated with viral particles after column purification, suggesting that Rep proteins bound to capsid may play a role in the early stages of AAV infection after entry. To test the role of CPSF6 in AAV infection, we performed replication experiments in CPSF6-depleted cells, which surprisingly revealed an increase in AAV replication in the context of CPSF6 knockdown. It is difficult to draw any conclusions from this experiment however as it was performed only once, and the general replication efficiency was quite poor. In addition, using replication efficiency as a readout for nuclear trafficking assumes that enhanced trafficking will necessarily and directly translate to improved replication. Given the multitude of events that occur after nuclear localization has already taken place, this is unlikely to be the case. For example, binding of CPSF6 to the capsid might be necessary for efficient trafficking while simultaneously slowing down the process of viral genome release, thereby rendering it almost impossible to tease these factors apart using replication as a readout. The better experiment would have been to analyze nuclear trafficking of the different AAV serotypes with and without CPSF6 using immunofluorescence.

In this chapter, we identified numerous putative interaction partners for the AAV Rep proteins using the screening method BioID in an effort to further elucidate the mechanism of Rep-mediated repression of heterologous

promoters. Despite certain limitations, our screen provided us with a coherent list of cellular factors involved in processes such as RNA metabolism, transcriptional regulation, chromatin remodeling, and the cellular DDR. Based on their respective cellular functions, we systematically assessed the role of several of these candidates in the regulation of the AAV p5 promoter by Rep. Although we were unable to demonstrate a role for these various candidates in repression, there is some evidence to suggest that TCERG1 and RUVBL1 may be necessary for efficient transcription, a possibility that requires further investigation. Moving forward, it will be necessary not only to test further candidates, but also to widen the scope of our functional validation such as to include roles outside of transcriptional repression. One possibility is to begin by assessing the replication efficiency of AAV in the presence or absence of candidate proteins, as this may provide the best end-point readout for the functional relevance of a protein-protein interaction in the AAV life cycle.

Chapter 4. KAP1 binds and represses latent AAV2 genomes

4.1 Introduction

Also identified in our BioID screens as an interaction partner for all four of the Rep proteins was the Kruppel-associated box domain-associated protein 1 (KAP1/TRIM28/TIF1β), a transcriptional corepressor comprised of an N-terminal RBCC protein-interaction domain, a central PxVxL region necessary for binding heterochromatin protein 1 (HP1), and a C-terminal plant homeodomain (PHD) and bromodomain, which recruit components of the NuRD histone deacetylase complex and the histone methyltransferase SETDB1 (Schultz et al., 2002; Schultz et al., 2001). KAP1 itself is recruited to the genome through interactions between its RBCC domain and KRAB-domain containing zinc fingers (KRAB-ZNFs) (Friedman et al., 1996), after which its PHD, bromodomain, and PxVxL region act cooperatively to form transcriptionally repressive heterochromatin characterized by the deposition of H3K9me3. This mechanism of KAP1-mediated repression has previously been linked to the silencing of viral elements such as MLV and ERVs (Rowe et al., 2010, 2013; Wolf and Goff, 2007). Additionally, KAP1 has been shown to be involved in the maintenance of CMV and KSHV latency (Chang et al., 2009; Rauwel et al., 2015; Sun et al., 2014). The observed interaction between Rep and KAP1 therefore provoked the question of whether KAP1 might also regulate AAV latency through epigenetic modification of AAV chromatin structure.

Both the latent and a lytic phase of the AAV life cycle require a distinct transcriptional program. It is already known for various other biphasic DNA viruses that the transition between programs is achieved through epigenetic modulation of chromatin dynamics (Knipe et al., 2013; Lieberman, 2008). Little is known however with respect to the nature and role of epigenetic marks in AAV infection and their contribution to viral genome organization and the temporal regulation of gene expression. Our understanding of AAV latency is further complicated by a degree of uncertainty regarding the relevance of integration *in vivo*. Although AAV sequences have been readily detected in a wide variety of human tissues (Chen et al., 2005; Friedman-Einat et al., 1997; Grossman et al., 1992; Han et al., 1996; Tezak et al.,

2000), evidence for integration comes primarily from observations in tissue culture, and so it is still unclear whether AAV establishes latency through integration or episomal persistence. There is some evidence to suggest that AAV assumes a chromatinized configuration shortly after infection, which would indicate a role for epigenetic modifications in the maintenance of latency and/or transcriptional regulation of AAV. An early study by CJ Marcus-Sekura and BJ Carter (Marcus-Sekura and Carter, 1983) used micrococcal nuclease digestion to demonstrate that intracellular AAV DNA was present in nucleosome-like structures similar to cellular nucleosomes. Surprisingly, this was true even in the absence of helper virus or AAV replication, which they hypothesized resulted from either the reassociation of infecting parental ssDNA, or by covalent integration into the genome. A later study investigating the persistence of rAAV vectors obtained from primate muscle up to 22 months after injection showed that rAAV also assumed a typical nucleosomal pattern, in this case however residing predominantly as monomeric and concatameric episomes that assimilated into chromatin (Penaud-Budloo et al., 2008). In the context of integration, genome-wide analysis of AAV integration sites has shown that, in addition to *AAVS1,* AAV integrates at numerous other genomic sites containing the requisite RBS with a preference for regions of open chromatin, further supporting a role for chromatin organization in AAV gene expression (Hüser et al., 2014; Petri et al., 2015).

Most viruses depend on specialized viral core proteins and/or cellular chromatin modifying proteins for the packaging of their genomes and transcriptional regulation. Adenoviral genomes are packaged as highly dense chromatin-like structures that must first undergo decondensation during nuclear entry in order to become replication competent (Giberson et al., 2012). Several cellular proteins contribute to this process, including the histone chaperone TAF1/Set, the ANP32A/B proteins, and the nucleolar protein nucleophosmin (Haruki et al., 2006; Matsumoto et al., 1993; Okuwaki et al., 2001). Similarly, HSV interacts with TAF1/Set to assemble an active chromatin structure on viral genomes (van Leeuwen et al., 2003), and the HSV VP16 protein recruits a wide range of histone acetyltransferases and chromatin-remodelling factors to regulate the temporal progression of latent

to lytic gene expression (Narayanan et al., 2007). These examples underscore the importance of chromatin dynamics, particularly in the life cycle of DNA viruses. It would therefore be interesting to determine if and what role these types of interactions might play in the regulation of AAV latency and lytic replication. Indeed, it has already been shown that AAV replication is also dependent on TAF1/Set activity (Pegoraro et al., 2006), which would suggest that AAV DNA can exist in a highly condensed form, which must be relaxed in order to replicate. Considering the influence of TAF1/Set on two of AAV's most common helper viruses, Ad and HSV-1, this is not entirely surprising and may in fact reflect another helper function of these viruses.

A better understanding of the relationship between AAV and epigenetic modification might also help to answer some crucial questions that remain regarding its biology. One phenomenon of AAV transcription that has yet to be mechanistically addressed is how the Rep proteins and different RBS can act as both activators and repressors of transcription. One possibility is that the formation of nucleosomes provides a scaffold upon which epigenetic modifications can act to rapidly "switch" a particular binding site from repressive to activating, and vice versa, by modifying the chromatin structure. Also, whether viewed as an evolutionary choice or merely as a replicative deficiency, helper virus dependency allows AAV to effectively defer gene expression until optimal conditions for replication are met, and it's possible that the definition of the contributing processes would provide insight into a novel form of latency in which epigenetic regulation is dependent on both helper and host.

4.2 Results

4.2.1 Transcriptional repression by Rep52 is not dependent on KAP1.

Samples from the BirA*-Rep52 screen were first analyzed by western blotting to confirm the presence of endogenous KAP1 among the purified proteins, revealing a band of approximately 110 kDa corresponding to KAP1 (Figure 20A). A repeat of the BioID in the context of exogenous FLAG-tagged KAP1 (FLAG-KAP1) and BirA*-Rep52 again confirmed the biotinylation of FLAG-KAP1 (Figure 20B). To gain further insight into the nature of the interaction between Rep and KAP1, FLAG-tagged proteins were co-IPed from lysates of 293T cells expressing FLAG-KAP1, or a FLAG-GFP control, with each of the Rep proteins. Conventional co-IPs were unsuccessful, however by using CL-IP we were able to detect interactions between KAP1 and at least three of the Rep isoforms (Figure 20C). Rep40 and Rep52 were highly abundant in the FLAG-KAP1 pull-down fraction, and Rep78 was also detectable despite having low expression levels in the input. A band corresponding to Rep68 was visible with both FLAG-KAP1 and FLAG-GFP (Figure 20C), indicating this may be a nonspecific interaction. This was confirmed in the reciprocal CL-IP, which validated KAP1 interaction with all Rep proteins except Rep68 (Figure 20D). As Rep40 represents the only shared domain between the Rep proteins, these results suggest that the ATPase domain is sufficient to mediate a Rep-KAP1 interaction. The requirement for cross-linking however would indicate that this interaction is transient, weak, and/or indirect.

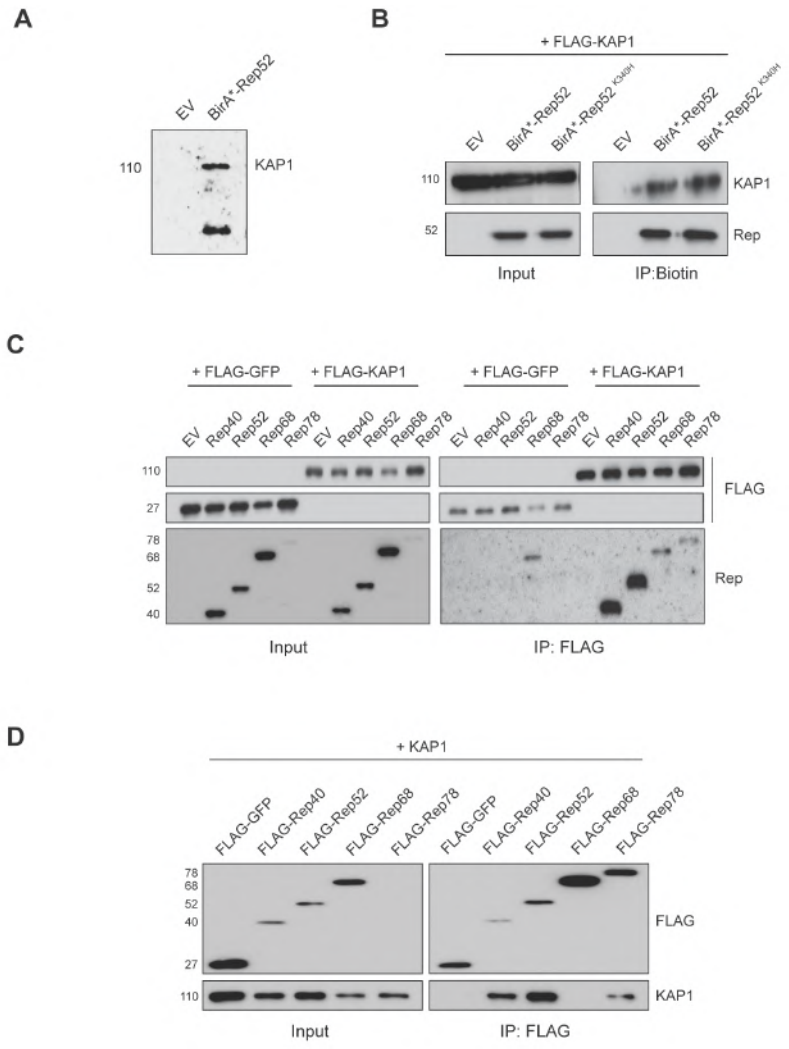


Figure 20. Validation of the physical association between Rep proteins and KAP1.
(A) Biotinylated proteins purified from the BirA*-Rep52 BioID screen, or the empty vector (EV) control, were analyzed by western blotting for the presence of KAP1. **(B)** Biotinylated proteins purified from cells expressing FLAG-KAP1 and BirA*-Rep52, BirA*-Rep52^{K340H}, or an empty vector (EV) control in the presence of free biotin were analyzed by western blotting for the presence of FLAG-KAP1 and Rep. **(C)** Cross-linked IP for FLAG-tagged proteins from lysates of 293T cells expressing FLAG-KAP1 or a FLAG-GFP control and each of the four Rep proteins. **(D)** Cross-linked IP for FLAG-tagged proteins from lysates of 293T cells expressing FLAG-GFP, FLAG-Rep40, FLAG-Rep52, FLAG-Rep68, or FLAG-Rep78 and KAP1 from transfected 293T cells.

To explore the possible significance of this interaction for Rep-mediated transcriptional regulation, we assessed the ability of Rep52 to repress the AAV2 p5 promoter in the context of KAP1 depletion. 293T cells were transduced with lentivectors targeting the 3' UTR of KAP1 (shKAP1), or the empty vector as a control (shEMPTY). Untreated cells were processed in parallel as a further control for off-target effects of the lentivector

transductions. 48 h after transduction, cells were transfected with the previously described p5-mCherry reporter construct alone, or with increasing concentrations of Rep52. 48h after transfection, cells were harvested for protein extraction, and p5 activity was assessed by western blotting for mCherry (Figure 21). As expected, expression of Rep52 in both untreated cells and control cells resulted in the efficient dose-dependent repression of p5 as demonstrated by a loss of mCherry expression. Disappointingly, Rep-mediated p5 repression was completely unaffected by depletion of KAP1. Taken together, these data indicate that a consistent physical association exists between the Rep proteins and KAP1, but that this interaction is not involved in transcriptional regulation by the Rep proteins, as measured by this assay. This does not exclude the possibility that this interaction bears functional significance in other aspects of the AAV2 life cycle however.

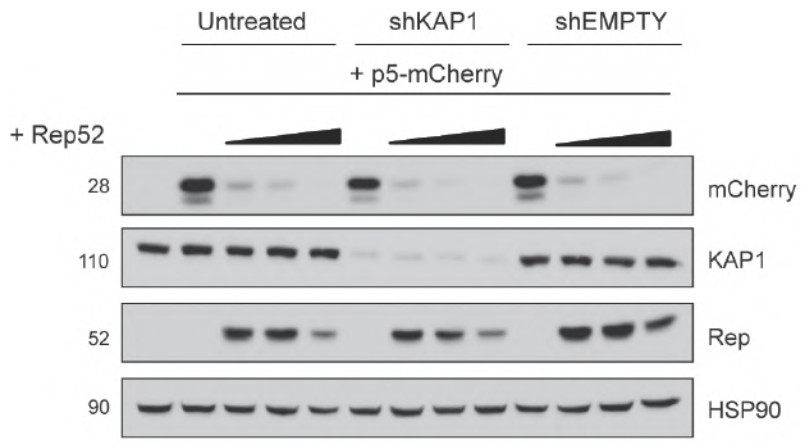


Figure 21. KAP1 is not necessary for Rep-mediated transcriptional regulation of the p5 promoter.
Untreated 293T cells or cells transduced with a lentiviral vector expressing a shRNA targeting the 3' UTR of KAP1 (shKAP1) or an empty vector control (shEMPTY) were transfected with a p5-mCherry reporter construct and increasing concentrations of Rep52 to determine the effect of KAP1 knockdown on Rep-mediated repression of the p5 promoter. Cells were harvested 48h after transfection, and p5 activity was determined by western blotting for mCherry.

4.2.2 Depletion of KAP1 leads to enhanced viral transcription and replication

To explore the possible significance of the Rep-KAP1 interaction in the AAV2 life cycle, we performed viral genome replication experiments in cells depleted for KAP1, as previously described for CPSF6. To determine the optimal infection conditions for these experiments, 293T cells infected with 10 - 10^4 IU/cell of AAV2, with or without Ad5 coinfection (2 PFU/cell), were harvested ~42h after infection for qPCR analysis of AAV2 replication (Figure 22A). In the absence of Ad5 coinfection, we observed a progressive increase in viral genomes with increasing MOIs. Interestingly however, Ad5 coinfection led to similar AAV2 replication levels in cells infected with 10, 100, or 1000 IU/cell of AAV2, suggesting a saturation point for AAV2 replication. Furthermore, cells infected at the highest AAV2 MOI showed ~10-fold lower replication levels, indicating that high-number input of viral genomes hinders rather than enhances replication, potentially by interfering with Ad5 replication too early in the lytic cycle and reducing the potential for AAV2 replication. Replication efficiency was determined by the normalization of each coinfected condition to its respective AAV2 input control (AAV2/-Ad5). This revealed that infection with 10 IU/cell of AAV2 resulted in the greatest dynamic range of replication (Figure 22B), providing us the best opportunity to detect changes in replication efficiency. Based on these results, we performed all subsequent infections using 10 IU/cell of AAV2 and 2 PFU/cell of Ad5, unless otherwise stated. Optimal concentration of Ad5 for 293T cells was previously determined in the lab based on cytopathic effect (CPE) 48h after infection.

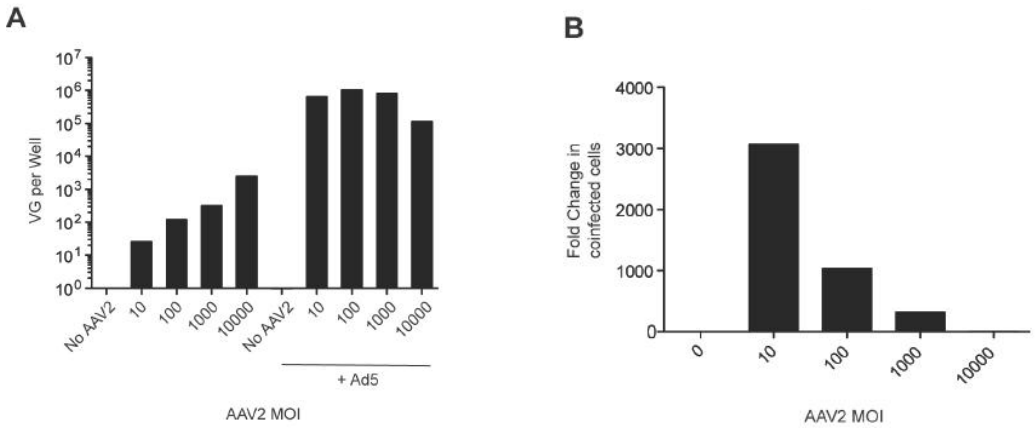


Figure 22. Determination of optimal infection conditions for genome replication experiments.
(A-B) 293T cells infected with 10, 100, 1000, or 10,000 IU/cell of AAV2, with or without Ad5 infection (2 PFU/cell), were harvested 48 h after infection for qPCR analysis of viral genomes. **(A)** Viral genomes per well. **(B)** Replication efficiency calculated by normalizing each coinfected condition to its respective input control (AAV2 alone).

For the genome replication experiments, control and KAP1-depleted 293T cells were infected with Ad5, AAV2, or Ad5+AAV2 in order to initiate productive replication, and were harvested for DNA, RNA, and protein extraction approximately 42h after infection with Ad5, or when cells began to display optimal cytopathic effect (CPE). Optimal CPE is defined by cells that display a rounded and enlarged phenotype, as opposed to the normal "star-shaped" morphology of HEK293T cells, and which are beginning to detach but still appear bright and healthy. Cells that had completely lifted by the time of harvest were deemed too advanced in the infection cycle and were excluded from analysis. In the absence of Ad5, AAV2 replication was undetectable by qPCR in both KAP1-depleted and control cells; this was not surprising as AAV2 is dependent on several helper factors to initiate replication. In the context of coinfection however, a 6-fold enhancement in AAV2 replication was observed in KAP1-depleted cells as compared to control cells (Figure 23A). Interestingly, we also observed a 2-fold enhancement in Ad5 replication in KAP1-depleted cells. Supporting the genome replication data, reverse transciptase qPCR (RT-qPCR) analysis of *rep* and *cap* transcripts in coinfected cells revealed a 6-fold enhancement in transcription of both viral genes in KAP1-depleted cells (Figure 23B). Rep and capsid protein levels were also elevated, and quantification of VP3 levels revealed a 5-fold

enrichment (Figure 23C). Rep protein levels were not quantified as the temporal regulation of Rep expression throughout the replication cycle results in protein levels that do not necessarily mirror replication efficiency.

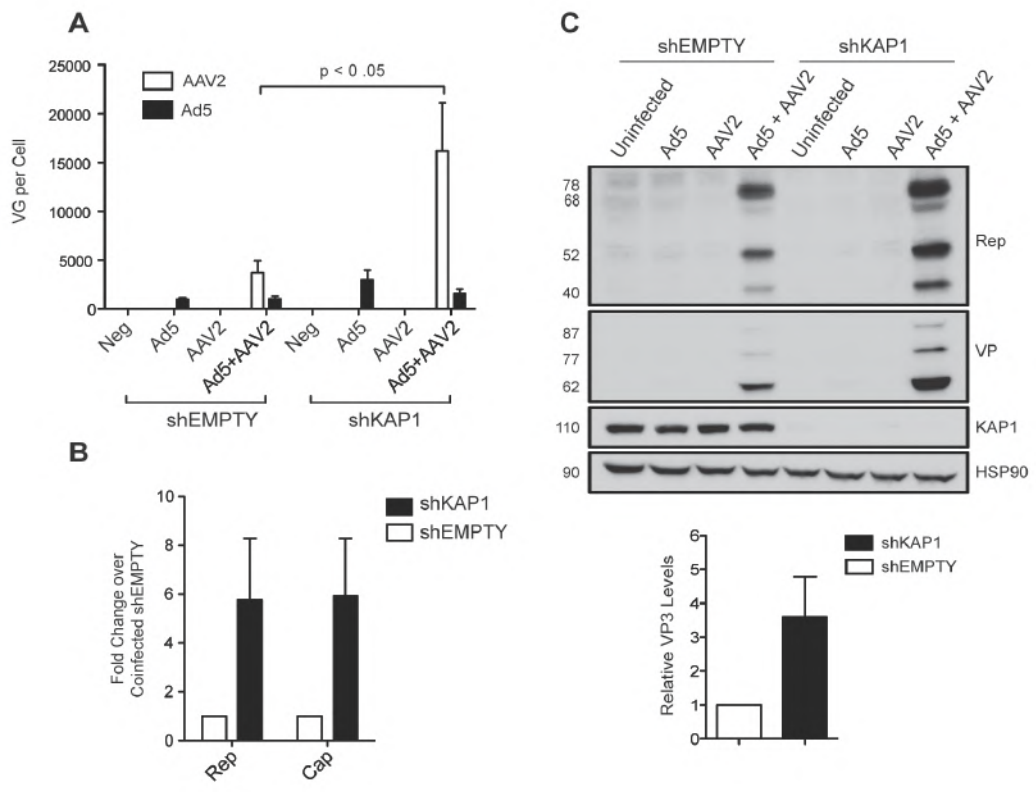


Figure 23. Knockdown of KAP1 leads to enhanced AAV2 replication and transcription
(A-C) Control (shEMPTY) or KAP1-depleted (shKAP1) 293T cells infected with either AAV2 (10 IU/cell) alone, Ad5 (2 PFU/cell) alone, or coinfected were harvested for DNA, RNA, and protein extraction ~42h after infection. **(A)** Viral genome replication analyzed by real time qPCR. **(B)** *Rep* and *cap* expression analyzed by RT-qPCR. Expression levels were calculated by the ΔΔct method and represent fold changes over cells infected with AAV2 alone. **(C)** Viral protein expression and KAP1 knockdown efficiency analyzed by western blotting using anti-Rep, anti-VP, and anti-KAP1 antibodies (upper panel). VP3 levels were quantified using ImageJ software (lower panel).
Data are reported as mean±SEM, n=5, and statistical significance was determined by unpaired t test.

We next repeated these experiments in HeLa cells to determine if this phenotype was specific to 293T cells. Efficiency of KAP1 depletion is shown in Figure 24A. A very modest 1.5-fold increase in AAV2 replication was observed in KAP1-depleted HeLa cells as compared to controls (Figure 24B). Ad5 replication in these cells however was vastly more efficient than in 293T

and completely eclipsed that of AAV2, making it difficult to conclude with any certainty whether the replication effect observed in 293T cells also extends to HeLa cells.

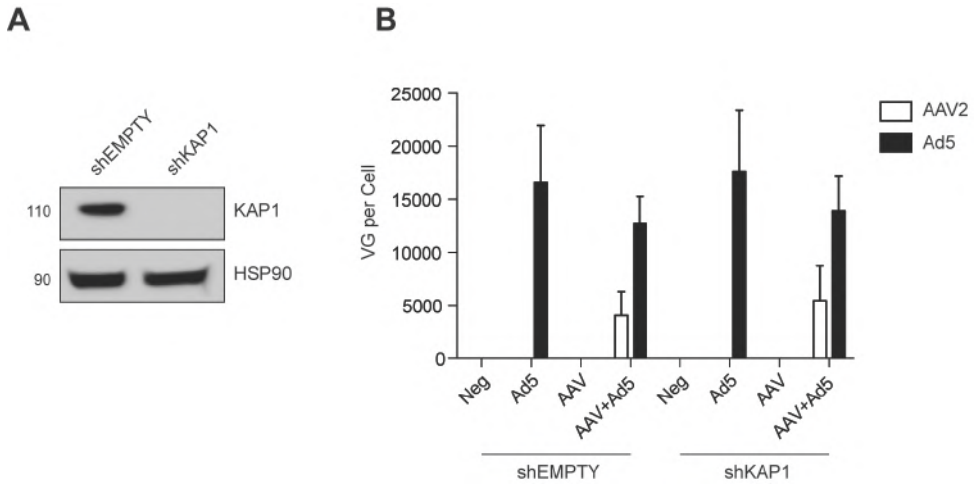


Figure 24. Genome replication in KAP1-depleted HeLa cells.
(A-B) Control (shEMPTY) and KAP1-depleted (shKAP1) HeLa cells were infected with AAV2 (10 IU/cell), Ad5 (2 PFU/cell), or coinfected with AAV and Ad5 and were harvested for DNA and protein extraction ~42h after infection **(A)** Western blot analysis of KAP1 knockdown. **(B)**. Viral genome replication analysed by real time qPCR.
Data are reported as mean±SEM, n=3.

To ensure that our observations were specific to KAP1 depletion, we next sought to demonstrate a rescue of baseline AAV2 replication levels in KAP1-depleted cells complemented with exogenous FLAG-KAP1. This proved to be challenging however, and it became apparent that achieving physiological levels of exogenously expressed FLAG-KAP1 would be crucial. To this end, we assayed multiple strategies for complementation. In the first set of experiments, KAP1-depleted cells were transfected with different concentrations of a CMV-driven FLAG-KAP1 expression construct and were coinfected 24h later with AAV2 and Ad5. Physiological levels of KAP1 expression were not achieved at either DNA concentration tested here. Instead, exceedingly high levels of exogenous KAP1 were associated with enhanced AAV2 replication, both in control and KAP1-depleted cells (Figure 25A). We next generated stable 293T cell lines expressing FLAG-KAP1

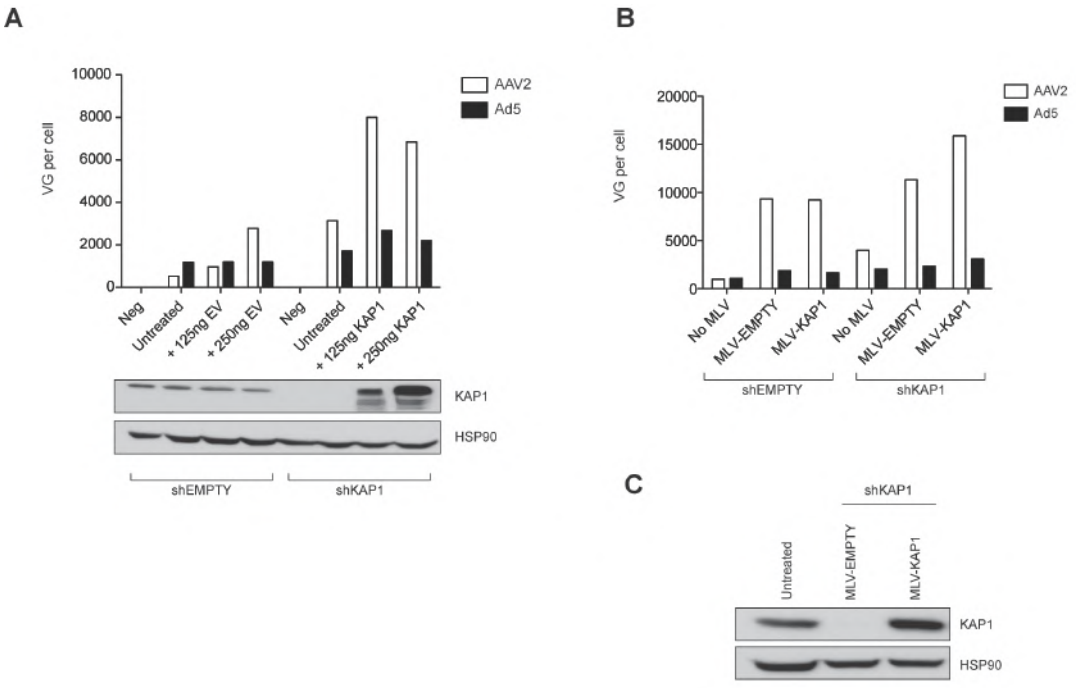


Figure 25. Viral replication in the context of KAP1 complementation.
(A) Viral replication in control (shEMPTy) or KAP1-depleted (shKAP1) 293T cells complemented with either empty vector (EV) or CMV-KAP1 (KAP1) and coinfected with AAV2 (10 IU/cell) and Ad5 (2 PFU/cell). KAP1 protein levels were monitored by western blotting. **(B)** Viral replication in control (shEMPTy) or KAP1-depleted (shKAP1) 293T cells stably expressing KAP1 (MLV-KAP1) or an empty expression cassette (MLV-EMPTY) and coinfected with AAV2 (10 IU/cell) and Ad5 (2 PFU/cell). shEMPTY and shKAP1 (No MLV) were infected in parallel to control for the effect of transduction with the retroviral expression vector. **(C)** KAP1 protein levels in untreated, shKAP1 + MLV-EMPTY, and shKAP1 + MLV-KAP1 293T cells.

using a retroviral expression vector (MLV-KAP1). These were depleted for endogenous KAP1 and coinfected 48h later with AAV2 and Ad5. While FLAG-KAP1 levels were much closer to physiological KAP1 levels here, AAV2 replication was elevated across all conditions (Figure 25B and C). Given that control cells (MLV-EMPTY + shEMPTY) also displayed greatly enhanced AAV2 replication as compared to cells treated with control shRNA alone (shEMPTY), we posited that integration of the retroviral vectors was affecting AAV2 replication. It is worth noting here that virtually any form of DNA damage has been shown to enhance AAV2 replication, and so it is conceivable that a cellular response to retroviral integration could be impeding our analysis.

Based on this, we returned to a transfection-based complementation strategy. FLAG-KAP1 was cloned into two different expression constructs in which truncated CMV promoters are used to achieve attenuated protein expression (Morita et al., 2012). These were tested for optimal expression levels (Figure 26A) and then used to complement KAP1-depleted cells as described in the first set of experiments. Levels of exogenous FLAG-KAP1 were comparable to physiological KAP1 levels here (Figure 26B), and complementation of KAP1-depleted cells with FLAG-KAP1 restored AAV2 replication levels to those observed in control cells transfected with an empty vector (EV) (Figure 26C). The observable rescue effect was modest however as transfection of the empty vector alone increased AAV2 replication in control cells, effectively eliminating the dynamic range of replication in which to demonstrate a rescue. Given that these data were not entirely conclusive, the replication phenotype was also recapitulated via siRNA-mediated knockdown of KAP1. Replication efficiency was generally lower, likely as a result of double siRNA transfection prior to infection, however we still observed a 2-fold enhancement in AAV2 replication in cells treated with siRNAs targeting KAP1 as compared to controls (Figure 26D). Taken together, these data indicate that the association between Rep and KAP1 bears functional significance in the AAV2 life cycle, and that depletion of KAP1 leads to an enhancement in transcription and genome replication by an unknown mechanism.

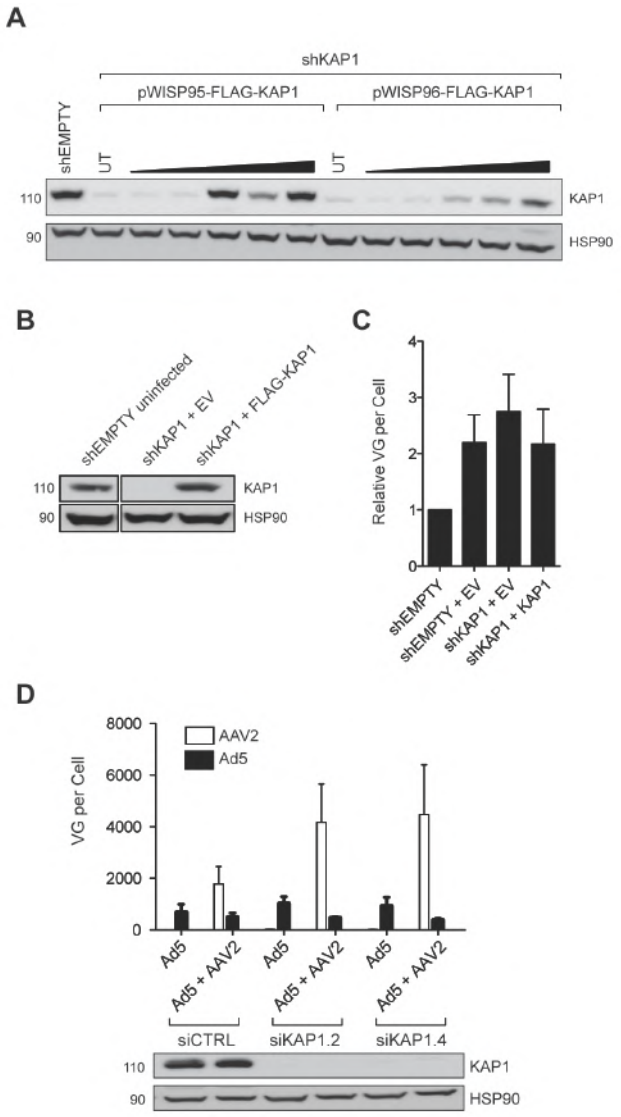


Figure 26. Validation of replication phenotype observed in KAP1-depleted 293T cells. **(A)** KAP1-depleted (shKAP1) 293T cells transfected with increasing concentrations of either pWISP95-FLAG-KAP1 or pWISP96-FLAG-KAP1, in which truncated CMV promoters are used to achieve attenuated levels of protein expression, were harvested 48 h after transfection. Levels of FLAG-KAP1 were determined by western blotting using anti-KAP1 antibody. Control cells (shEMPTY) were used for comparison. **(B, C)** Control and KAP1-depleted 293T cells that had been complemented with either an empty vector control or pWISP95-FLAG-KAP1 were coinfected with AAV2 (10 IU/cell) and Ad5 (2 PFU/cell). Cells were harvested 48 h after infection for DNA and protein extraction. **(B)** Analysis of KAP1 protein levels by western blotting using anti-KAP1 antibody. **(C)** Real time qPCR analysis of viral genomes. Data are normalized to replication levels in control cells (shEMPTY) and are reported as mean±SEM, n=3. **(D)** 293T cells transfected with two siRNAs targeting KAP1 (siKAP1.2 and siKAP1.4), or a non-targeting control (siCTRL) were infected with Ad5 (2 PFU/cell), or coinfected with AAV2 (10 IU/cell) and Ad5 (2 PFU/cell) and harvested 48 h after infection for qPCR analysis of viral genome, and western blot analysis to determine KAP1 knockdown efficiency. Data are reported as mean±SEM, n=4.

4.2.3 KAP1 binds the AAV2 genome during latency resulting in trimethylation of AAV2-associated H3K9.

Given that the association of KAP1 with the genomes of CMV and KSHV is known to repress lytic gene expression and maintain latency, we next asked if KAP1 could be repressing AAV2 through the binding of its genome and subsequent formation of heterochromatin. We performed KAP1-specific chromatin immunoprecipitation (ChIP) experiments on control and KAP1-depleted 293T cells that were infected with AAV2 alone in order to produce a latent infection. IgG was used as a control. Chromatin was isolated 48h after infection and analyzed by qPCR using primers specific for various regions of the AAV2 genome, as well as GAPDH as a negative control, and two zinc finger genes, ZNF180 and ZNF274, as positive controls for KAP1 binding. KAP1 binding was detected across *rep*, particularly at the 5' and middle regions, and this binding was accordingly lost in KAP1-depleted cells confirming the observed signal was in fact KAP1-dependent (Figure 27A). To determine the functional significance of this binding, we also performed ChIP for trimethylated H3K9 (H3K9me3), a known marker for KAP1-mediated repression. H3K9me3 was detected across the AAV2 genome (Figure 27B), and, importantly, histone methylation was lost in KAP1-depleted cells at a ratio similar to that observed for ZNF180 and ZNF274, supporting the notion that KAP1 not only binds the AAV2 genome but does so in the capacity of a repressor. Interestingly, no specific binding of KAP1 or enrichment for H3K9me3 was detected at the viral p5 promoter, a region known to be regulated by Rep and necessary for the transactivation of all three viral promoters. Indeed, we demonstrated earlier that the ability of Rep to regulate expression from a p5-mCherry reporter construct was completely unaffected by depletion of KAP1 (Figure 21).

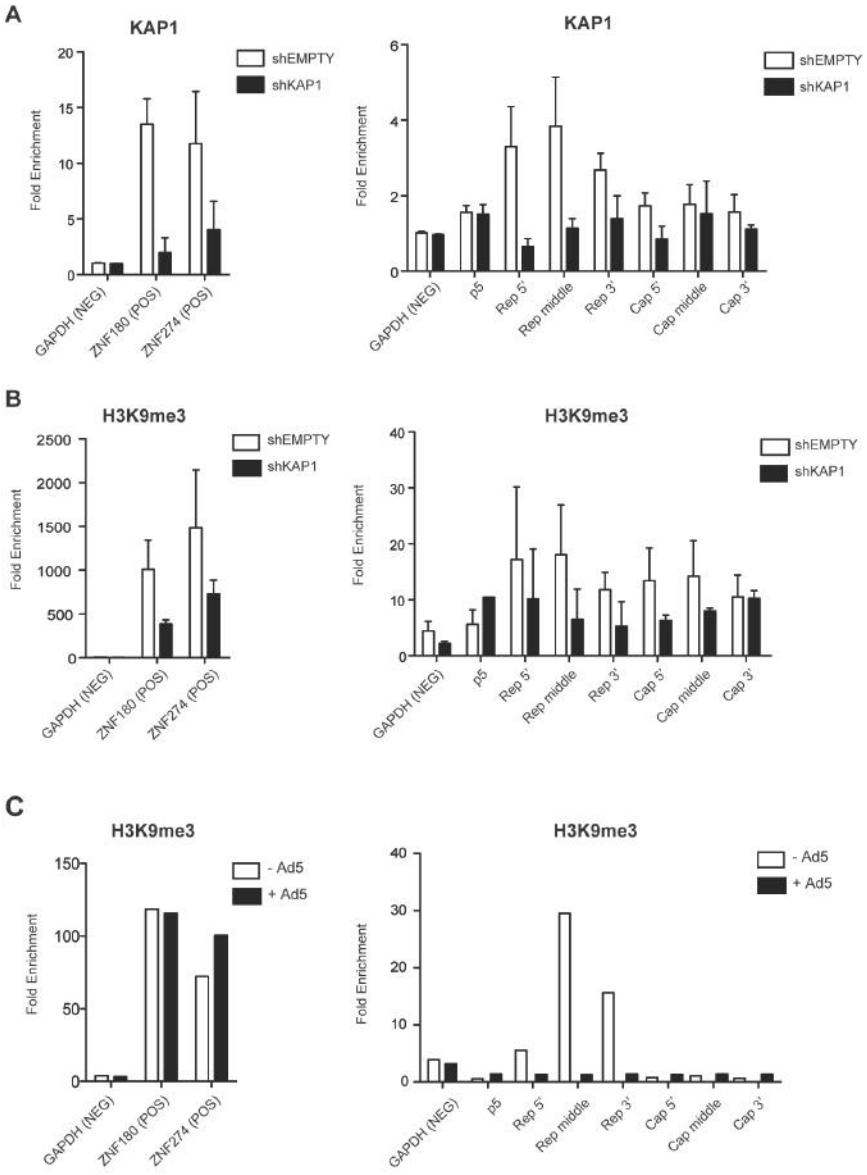


Figure 27. The latent AAV2 genome is methylated by recruitment of KAP1 to the *rep* ORF

(A) ChIP-qPCR was performed on control (shEMPTY) and KAP1 depleted (shKAP1) 293T cells 48h after infection with AAV2 (100 IU/cell) alone, using anti-KAP1 antibody or IgG. Purified chromatin was analyzed by qPCR using primers specific for the viral p5 promoter or various regions of the *rep* and *cap* ORFs (right panel). *GAPDH* was used as a negative control, and the zinc finger genes *ZNF180* and *ZNF274* were used as positive controls (left panel). Values are reported as mean±SEM for 3 independent experiments. **(B)** ChIP-qPCR was performed as described above, using anti-H3K9me3 antibody and IgG. Values are reported as mean±SEM for 3 independent experiments. **(C)** ChIP-qPCR was performed 48h after infection as described above on 293T cells that were infected with AAV2 (100 IU/cell) alone, or coinfected with AAV2 and Ad5 (2 PFU/cell).

To determine whether productive AAV2 replication was associated with a loss in trimethylation of AAV2-associated H3K9, we performed parallel H3K9me3-specific ChIP in 293T cells infected either with AAV2 alone, or coinfected with AAV2 and Ad5 to induce lytic replication. Chromatin was isolated 48 h after infection and analyzed by qPCR as described above. In the context of latent AAV2 infection, enrichment for H3K9me3 was again apparent across *rep* (Figure 27C). In contrast to what we observed in the previous set of experiments however, H3K9me3 enrichment was detectable predominantly at the middle to 3' region of *rep,* rather than the 5' to middle region. Surprisingly, there was also a lack of any observable enrichment for H3K9me3 across *cap* (Figure 27C). In agreement with the previous experiments, the p5 promoter was again free from the repressive mark. Importantly, no AAV2-associated enrichment for H3K9me3 was detected in cells coinfected with AAV2 and Ad5 (Figure 27C), suggesting that replication is associated with a loss of KAP1 repressor activity. It is important to note however that the interpretation of these data is complicated by the presence of efficient AAV2 genome replication in coinfected cells. It is possible that high levels of AAV2 DNA in the input samples of coinfected cells could mask a potential signal for H3K9me3 enrichment, resulting in a false negative. For this reason, we performed this experiment only once. This might help to explain not only the difference we observed in the pattern of *rep* and *cap* histone methylation, but also the apparent lower degree of H3K9me3 enrichment at the positive controls than previously observed (Figure 27B), as a degree of variability was inherent to these experiments.

4.2.4 KAP1 recruits CHD3 and SETDB1 to mediate H3K9 methylation of the latent AAV2 genome.

We then reasoned that if KAP1 represses AAV2 through the canonical pathway involving recruitment of SETDB1 and the NuRD complex, depletion of other members of the repressive complex would have a similar effect on AAV2 replication as depletion of KAP1. To test this, 293T cells were transfected with siRNAs targeting either CHD3 or SETDB1 and then infected as described above. AAV2 replication was monitored both by western blot for the VP proteins, and by qPCR. A 6-fold increase in VP expression was

observed in cells depleted for either CHD3 or SETDB1, and the depletion of both appeared to have a cumulative effect, resulting in a 10-fold increase in VP levels as compared to control cells (Figure 28A and B). The effect was less pronounced at the level of genome replication, although the trend remained the same. Cells depleted for either CHD3 or SETDB1 demonstrated a 2-fold enhancement in AAV2 replication, and cells depleted for both a 3-fold enhancement (Figure 28C). Taken together, these data strongly suggest that KAP1 is repressing AAV2 through the binding of AAV2 *rep* and the subsequent recruitment of histone and chromatin modifying proteins, which then act together to methylate AAV2-associated H3K9.

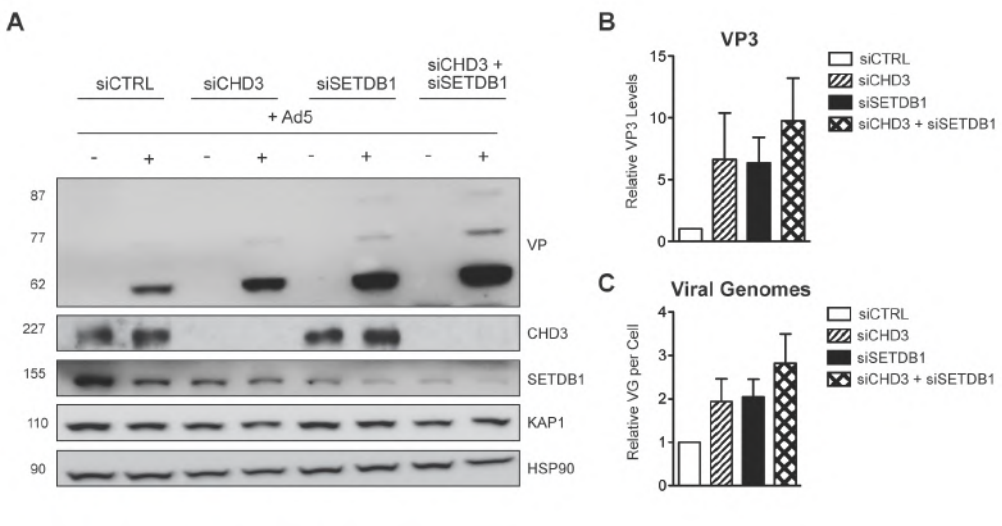


Figure 28. Depletion of CHD3 and SETDB1 leads to enhanced AAV2 replication and protein expression.

(A-C) 293T cells were transfected with non-targeting control siRNA or siRNA targeting CHD3 or SETDB1, and were infected with Ad5 (2 PFU/cell) alone, or coinfected with Ad5 and AAV2 (10 IU/cell). Cells were harvested for DNA and protein extraction. **(A)** VP protein expression and depletion of CHD3 and SETDB1 were analysed by western blotting using anti-VP, anti-CHD3, and anti-SETDB1 antibodies **(B)** Quantification of VP3 levels using ImageJ software **(C)** Viral genome replication analyzed by real time qPCR. Values are reported as mean±SEM, n=4.

4.3 Discussion

It is becoming increasingly clear that the assembly, modification, and remodeling of viral chromatin play a crucial role in determining the outcome of infection. DNA viruses in particular capitalize on the formation of nucleosomes in order to establish a dynamic state in which transcriptional activity can be rapidly altered through the activity of cellular histone- and chromatin- modifying factors (Knipe et al., 2013; Lieberman, 2008). Indeed, the fates of the AAV helper viruses Ad5 and HSV-1 are inextricably linked to their chromatin states (Arbuckle and Kristie, 2014; Knipe and Cliffe, 2008). In addition to its role in the temporal regulation of viral gene expression, the formation of repressive chromatin on viral genomes can also be the result of a cellular defence strategy, as has also been observed for Ad, HSV-1, and hepatitis B virus (HBV) (Ducroux et al., 2014; Schreiner et al., 2013a, 2013b). There is still much to learn with respect to the mechanisms involved in the temporal regulation AAV transcription; in particular, (1) what factors are necessary for the establishment and maintenance of latency and what role do these play in supporting/suppressing reactivation, (2) what changes occur to allow transition from latency to lytic replication, and (3) how might these changes evolve throughout the course of infection. In this chapter, we first show that the AAV2 Rep proteins and transcriptional corepressor KAP1 physically interact. We also demonstrate that recruitment of KAP1 to the Rep ORF leads to the trimethylation of AAV2-associated H3K9, and we provide evidence that KAP1 achieves this through its canonical mechanism of repression involving the recruitment of the nucleosome- and histone-modifying proteins, CHD3 and SETDB1. Depletion of KAP1 led to a significant enhancement in AAV2 replication and transcription, further supporting the notion that KAP1 acts as a repressor of AAV2.

Given how intimately intertwined the fate of AAV is with that of both Ad5 and HSV-1, it is difficult to envision an AAV life cycle in which chromatin dynamics are not involved. Here, we demonstrate a role for KAP1 in the epigenetic regulation of the AAV2 genome. Knockdown of KAP1 resulted in a significant enhancement in viral DNA replication, transcription, and protein expression, suggesting that AAV2 is a target of KAP1 activity. Not only does this provide evidence for the chromatinization and histone

methylation of AAV2, it also predicts a requirement for the remodeling of viral chromatin in overcoming KAP1-mediated repression during the transition from latency to lytic replication. Several proteins pulled down in our original BioID screen support the idea that AAV becomes chromatinized upon uncoating. Interactions were detected between Rep52 and CHAF1-A, a core component of the CAF-1 complex thought to mediate chromatin assembly in DNA replication and repair (Moggs et al., 2000; Smith and Stillman, 1989), and which is required for chromatin assembly during SV40 DNA replication *in vitro* (Stillman, 1986). Also identified were TOX4 and PPP1R10, two proteins that together form the PTW/PP1 complex implicated in the regulation of chromatin structure during mitosis (Lee et al., 2010). Even in the context of latency, the AAV genome must undergo some minimal level of replication during cellular S phase in order to generate the dsDNA construct necessary for long-term persistence and/or integration. It is thus possible that the CAF-1 and PTW/PP1 complexes mediate nucleosome formation during this process. The viral ITRs are also known to mimic DNA DSBs and could trigger DNA repair, leading to nucleosome formation (Moggs et al., 2000). In addition, TAF1-B/Set, nucleophosmin, and ANP32A were pulled down with at least one of the Rep proteins. Together, these three proteins comprise the cellular Template Activating Factor (TAF) activity necessary for the remodeling of incoming adenovirus DNA (Haruki et al., 2006; Xue et al., 2005), and it is tempting to speculate that they provide the same function for AAV. Similarly, the cellular factor host cell factor 1 (HCF-1) necessary for HSV-1 reactivation (Narayanan et al., 2007) was also identified. These interactions strongly suggest that AAV undergoes a similar transformation from transcriptionally repressed heterochromatin to a more relaxed and active chromatin state in preparation for replication. Interestingly, the chromatin boundary factor CTCF was also found to interact with Rep52 in our screen. This protein is known to insulate the HSV-1 LAT promoter and enhancers from the ICP0 and other immediate early gene promoters (Amelio et al., 2006), and has also been implicated in the regulation of Epstein Barr Virus (EBV) epigenetic gene regulation (Chau et al., 2006).

The manner in which AAV2 establishes latency is still unclear. Evidence for integration *in vitro* is compelling and has led to the detailed molecular characterization of AAV2 integration into *AAVS1* (Henckaerts and Linden, 2010). However, studies attempting to characterize AAV DNA from human tissues have found evidence mostly for episomes (Schnepp et al., 2005). Although we cannot exclude that some of the observed binding of KAP1 to the AAV2 genome occurred in the context of integrated viral DNA, we believe that the majority of interactions occur with episomal DNA considering the conditions used for infection in our ChIP experiments. Cells were infected with 100 IU/cell of AAV2 and were harvested 48 hours after infection without passaging. While 30-40% of clonal cell lines have been shown to contain integrated viral DNA at this MOI (Hamilton et al., 2004), it is likely that the high level of non-integrated virus present in the input sample would mask any enrichment if KAP1 were predominantly bound to integrated viral DNA. Rather, our data are consistent with earlier studies showing that AAV2 adopts a nucleosomal configuration shortly after infection (Marcus-Sekura and Carter, 1983), and further indicate that heterochromatinization by KAP1 may silence the viral genome without the need for integration.

In the absence of helper virus coinfection, AAV2 establishes latency and is silenced through the simultaneous binding of the p5 promoter region by the cellular factors YY1 and MLTF. To initiate lytic replication, AAV2 is dependent on the disruption of these repressive factors by the activity of the helper virus factors Ad5 E1A or HSV-1 ICP0. Here, we show that KAP1 provides a second layer of regulation, the antagonism of which is necessary but not sufficient for reactivation. Depletion of KAP1 resulted in a significant enhancement in AAV2 genome replication, transcription, and protein expression in cells coinfected with AAV2 and Ad5. This effect was not observed in cells infected with AAV2 alone however, reflecting the dependency of AAV2 on various helper factors to initiate replication. These observations bear interesting similarities to those made for KAP1 regulation of KSHV. Like AAV2, KSHV replication was enhanced, but not triggered, by KAP1 depletion, both in the context of induced KRta expression and hypoxia-induced KSHV reactivation (Chang et al., 2009; Zhang et al., 2014).

It is also interesting to note that KSHV is an episomal virus with the ability to modulate KAP1 activity via its viral kinase, vPk (Chang et al., 2009). Based on these observations and the fact that both AAV2 and KSHV are clearly capable of replication without the need for KAP1 depletion, it is tempting to speculate that these viruses have domesticated KAP1 repression to their advantage. The fact that KAP1 depletion does nevertheless affect AAV2 replication may simply reflect a difference in replication kinetics, with replication in KAP1-depleted cells being able to begin more rapidly. This would have a two-fold effect – (1) the more rapidly initiated replication would likely lead to better competition with Ad5, tipping the scales in favour of AAV, and (2) replication would simply have progressed further at the time of harvest. Alternatively, the presence of concurrent Ad5 replication might create a fixed window of opportunity for AAV replication, and a more rapid initiation would naturally translate to higher final numbers.

KAP1 has been shown to regulate the latency of both CMV and KSHV. KAP1 recruitment to the CMV genome leads to H3K9me3 deposition across various lytic genes, while latency-associated genes remain free from repressive marks (Rauwel et al., 2015). Similarly, LANA-mediated recruitment of KAP1 to the KSHV genome is involved in the shutdown of lytic gene expression during early stages of KSHV infection (Sun et al., 2014). Here, we observed KAP1-dependent H3K9me3 deposition across the latent AAV2 genome spreading downstream from KAP1 recruitment sites in *rep*. While AAV2 does not contain a true latency-associated gene, neither KAP1-binding nor H3K9me3 enrichment was detected at the p5 promoter, a region whose transactivating activity is necessary for initiation from all three viral promoters. It is possible that Rep and YY1 bound to p5 act not only to repress transcription during latency, but also to maintain the promoter free from repressive marks and competent for rapid reactivation. This could result either from direct competition between Rep and KAP1 recruitment factors for binding sites within p5, or, alternatively, Rep-dependent phosphorylation of KAP1 could promote the release of KAP1 from bound recruitment factors.

Two types of genomic KAP1 binding sites have been identified (Iyengar et al., 2011). Binding sites at the 3’ ends of ZNF genes resemble viral KAP1 binding sites in their enrichment for H3K9me3, while binding sites

near transcription start sites (TSS) are free from histone methylation and may instead be related to the recently discovered role for KAP1 in proximal promoter pausing (McNamara et al., 2016). KAP1-mediated recruitment of the 7SK snRNP complex and its associated pTEFb to sites of paused RNA pol II is thought to provide a constant source of pTEFb ready to phosphorylate and activate RNA pol II upon stimulation. While we show here that enrichment of AAV2-associated H3K9me3 is dependent on KAP1 recruitment, the proximity of KAP1 binding to the 5' region of *rep* is also analogous to binding sites near TSS. Further work will be necessary to fully elucidate the potential contribution of each mechanism.

It is unclear at this point how KAP1 is recruited to the AAV genome. The most well-characterized recruitment partners for KAP1 are the KRAB-ZFPs, although KRAB-only proteins as well as various other transcription factors are also known to interact with KAP1. Furthermore, mutant KAP1, in which the RBCC domain required for interaction with KRAB-ZFPs is removed, is still recruited to numerous sites in the genome (Iyengar et al., 2011). Although we show that KAP1 and the Rep proteins physically interact in a manner dependent on the Rep ATPase domain, it is unlikely that KAP1 is being recruited to the AAV2 genome by the Rep proteins themselves. The requirement for cross-linking of cells for detection of the Rep-KAP1 interaction suggests that Rep and KAP1 interact only indirectly. Also, as we did not DNase treat lysates prior to IP, we cannot be certain that the observed interactions were not a result of DNA bridging, although neither Rep40 nor Rep52 contain specific DNA binding activity. Furthermore, if Rep78 recruited KAP1 to the AAV2 genome, we would expect to observe KAP1 binding at the p5 promoter, which we do not. While we did not identify any KRAB-ZFPs in our BioID screens, we did identify several zinc finger-containing proteins, including ZNF318, a TF implicated in the transcriptional repression of androgen receptor (AR)-regulated genes during spermatogenesis (Ishizuka et al., 2005). Although very little is known about the function of ZNF318, it has been shown to interact with both MDM2 and HDAC2 (Okoro et al., 2013; Tao et al., 2006), two proteins that also interact closely with KAP1 (Schultz et al., 2001; Wang et al., 2005).

Based on the location of the primers used for ChIP-qPCR and the observed binding of KAP1 at the 5' to middle region of *rep*, it is quite possible that KAP1 is in fact binding to the p19 promoter. The p19 promoter contains two TATA sites, which recruit TBP, two Sp1 binding sites, and a site at position -140 that binds an unidentified cellular AAV-activating protein (cAAP) (Chejanovsky and Carter, 1989; Pereira and Muzyczka, 1997b; Srivastava et al., 1983). Transactivation of p19 by Rep appears to require both an Sp1 site and the cAAP site (Pereira and Muzyczka, 1997b) and is mediated by interactions between Sp1 bound at p19 and Rep bound at p5, which bring the p5 YY1 complex into proximity to p19 (Lackner and Muzyczka, 2002). Interestingly, recruitment of KAP1-7SK snRNP complexes to sites of paused RNA pol II exactly mirrors the occupancy profiles of Sp1 and TBP (McNamara et al., 2016), which are components of pre-initiation complexes and which both interact with Rep. It is conceivable then that KAP1-7SK snRNP complexes are bound to the p19 promoter, where they play a role in the release of paused RNA pol II upon stimulation by helper virus proteins. Furthermore, KAP1 has been shown to bind directly to YY1 in mouse ES cells (Schlesinger et al., 2013). Although this interaction was lost upon differentiation, it is possible that KAP1 and YY1 may still interact in other differentiated cells where both proteins are present. Interactions between KAP1 bound at p19 and YY1 bound at p5 may help to form the DNA loops necessary for transactivation. Along these lines, Ad E1A-activated YY1 may directly provide the cue for release of paused RNA pol II at p19, and potentially p40, through interactions with KAP1-7SK snRNP complexes. Arguing against this possibility however is the fact that depletion of KAP1 results in a decrease in the magnitude of induction of KAP1-7SK snRNP-regulated promoters (McNamara et al., 2016), whereas we observed enhanced transcription and replication in the context of KAP1 depletion. Furthermore, KAP1 binding to the AAV2 genome results in downstream histone methylation, which is not known to be characteristic of KAP1 binding sites associated with proximal promoter pausing. Further analysis of p19 transcription in the absence of helper virus coinfection will be necessary to determine the potential contribution of this mechanism to AAV gene regulation.

Many viral pathogens must contend with, modulate, and utilize host chromatin machinery during infection to promote efficient lytic replication, control persistent latent states, and circumvent potential cellular chromatin repression. In this chapter, we demonstrate for the first time a role for the epigenetic regulation of AAV replication and gene expression. We show that the latent AAV2 genome is chromatinized and repressed through KAP1-mediated histone methylation. These data place our budding understanding of AAV chromatin regulation on firmer footing alongside what is already known for other DNA viruses, and further presents the possibility that observations made for other viruses, in particular Ad5 and HSV-1, might also be extrapolated to AAV. In the following chapter, we investigate the mechanisms through which AAV is able to counteract this repression so as to support lytic replication.

Chapter 5. The AAV2 Rep proteins counteract KAP1 repression of latent viral genomes through antagonism of the protein phosphatase PP1

5.1 Introduction

In the previous chapter, we demonstrated that recruitment of KAP1 to the latent AAV2 genome leads to methylation of AAV2-associated H3K9. Given that AAV2 is capable of replication without the need for KAP1 depletion however, we next asked how the transcriptionally repressive activity of KAP1 is modulated during lytic AAV2 replication such as to enable viral gene expression and DNA replication. One possibility lays in the numerous post-translational modifications of KAP1 – in particular, phosphorylation of KAP1 at serine 824 (p-KAP1-S824). Upon induction of DNA DSBs, ATM-mediated phosphorylation of KAP1-S824 interferes with the interactions between KAP1, CHD3, and SETDB1, leading to release of the repressive complex and an associated global relaxation of heterochromatin (Goodarzi et al., 2011; Noon et al., 2010; Ryan et al., 1999; Ziv et al., 2006). This has been shown to be necessary for efficient DNA repair and enhanced survival after ionizing radiation, an effect thought to result from the increased access of DNA repair factors to sites of damage (White, 2006).

Phosphorylation of KAP1-S824 is also associated with a secondary loss of KAP1 corepressor activity. Indeed, phosphorylation of KAP1-S824 was shown to enhance *p21, Gadd45α, Bax, Puma*, and *Noxa* upregulation after DSB induction (Lee et al., 2007; Li et al., 2007b). Phosphorylation of KAP1-S824 is also linked to the transcriptional derepression of KAP1-regulated viral elements. During KSHV infection, KAP1 is recruited to the genome by the viral LANA protein where its repression of lytic genes, such as k-bZIP and vPK, is crucial to the maintenance of latency (Sun et al., 2014). Induction of lytic replication through the overexpression of the viral transcriptional regulator K-Rta results in the disassociation of KAP1 and a subsequent loss of HP1 and H3K9me3, and, consequently, reactivation of the viral genome (Chang et al., 2009). As reactivation progresses, the newly expressed viral kinase vPK inactivates KAP1 corepressor activity through phosphorylation of KAP1-S824, suggesting that counteraction by the virus is

necessary for full reactivation and maintenance of lytic replication (Chang et al., 2009). Similar results were observed for CMV, where p-KAP1-S824 was found to be associated with actively replicating but not latent viral DNA. The investigators here suggested that phosphorylation of KAP1-S824 was necessary to render the CMV genome permissive to transcription through the relaxation of viral heterochromatin, after which full viral gene expression became possible through NF-κB activation (Rauwel et al., 2015).

Upon completion of DNA repair, steady-state levels of KAP1-mediated heterochromatin and transcriptional repression are restored by the serine/threonine phosphatases PP4 and PP1 (Lee et al., 2012; Li et al., 2010), which serve to remove the phosphate group from KAP1-S824, thereby restoring KAP1 interactions with CHD3 and SETDB1. The PP1 catalytic subunit (PP1cs) exists in several different isoforms – PP1α, PP1β, or PP1γ – of which PP1α/β have been shown to regulate KAP1-S824 phosphorylation (Li et al., 2010). PP1 is involved in regulating a vast array of cellular activities, including glycogen metabolism, cell progression/mitosis, the cellular DDR, and protein synthesis (Ceulemans and Bollen, 2004; Cohen, 2002; Tang et al., 2008). This is achieved through the association of the PP1cs with numerous PP1 regulatory subunits, such as the nuclear inhibitor of PP1 (NIPP1), which are necessary to regulate the pleiotropic effects of PP1 activity by correctly targeting PP1cs to its various substrates. PP1cs interacts with its regulatory subunits through a consensus PP1-docking motif, $[KR][X]_{0-1}[VI]\{P\}[FW]$, the mutation of which often abolishes the binding between PP1cs and its regulatory components (Hendrickx et al., 2009).

In this chapter, we investigated the possible relationship between the phosphorylation of KAP1-S824 and AAV2 replication. Given the observed repression of the AAV2 genome by KAP1 under latent conditions, we hypothesized that phosphorylation of KAP1-S824 would be necessary to relieve KAP1 corepressor activity and allow for the initiation of AAV2 gene expression and replication. We found that levels of p-KAP1-S824 are greatly enhanced during wtAAV2 replication, and we demonstrate a correlation between the observed enhanced levels of p-KAP1-S824 and interactions

between the AAV2 Rep52 and Rep78 proteins, the phosphatase PP1, and the nuclear inhibitor of PP1 (NIPP1). Furthermore, we show that Rep proteins unable to support the interaction with PP1 and NIPP1 are also unable to support AAV2 transcription and replication, demonstrating a potential role for this interaction in the AAV2 life cycle.

5.2 Results

5.2.1 AAV2 replication correlates with the inactivation of KAP1 corepressor activity through phosphorylation of KAP1-S824

Upon the induction of DNA double strand breaks (DSB), ATM-dependent phosphorylation of KAP1 at serine 824 (S824) results in release of the repressive complex, relaxation of heterochromatin, and relief of transcriptional repression (Goodarzi et al., 2011; Li et al., 2007). We questioned whether AAV2 replication was associated with phosphorylation of KAP1-S824, which would suggest a requirement for the inactivation of KAP1 corepressor activity. We first explored this possibility by monitoring levels of phosphorylated KAP1-S824 (p-KAP1-S824) in cells infected with increasing MOIs of either AAV2 or recombinant AAV2 (rAAV2) in the presence of Ad5. rAAV2 is comprised of only the viral ITRs flanking a GFP transgene cassette, and as such is replication defective. However the ITRs are known to recruit components of the Mre11/Rad50/NBS1 (MRN) complex – the principal mediator of ATM activation – resulting in the silencing of rAAV genomes (Schwartz et al., 2007). While this mechanism appears to be independent of the activation of downstream effectors (Lentz et al., 2015; Schwartz et al., 2009), it has nevertheless been shown that $ATM^{-/-}$ cells display enhanced rAAV transduction (Cataldi and McCarty, 2013). We therefore used rAAV2 to control for the input of these structures. A clear dose-dependent increase in p-KAP1-S824 was observed in cells infected with AAV2, but not rAAV2 (Figure 29A). Ad5 infection alone did not trigger any observable phosphorylation of KAP1-S824. As this suggested a requirement for active replication, we next performed a time course to determine if phosphorylation of KAP1-S824 could be related to replication and coinciding with Rep expression. 293T cells were infected with either AAV2 or rAAV2 in the presence of Ad5, and p-KAP1-S824 levels were monitored by western blotting at 4h, 18h, 28h, and 42h post infection. In cells coinfected with AAV2 and Ad5, high levels of p-KAP1-S824 were apparent by 18 h post infection. Furthermore, the increase in p-KAP1-S824 levels correlated very well with the onset of Rep expression (Figure 29B). Phosphorylation of KAP1-S824 was also detectable in the context of Ad5 infection alone but appeared to be much weaker and occurred at later time points, pointing to the potential for

two independent pathways to the phosphorylation of KAP1-S824 by AAV2 and Ad5.

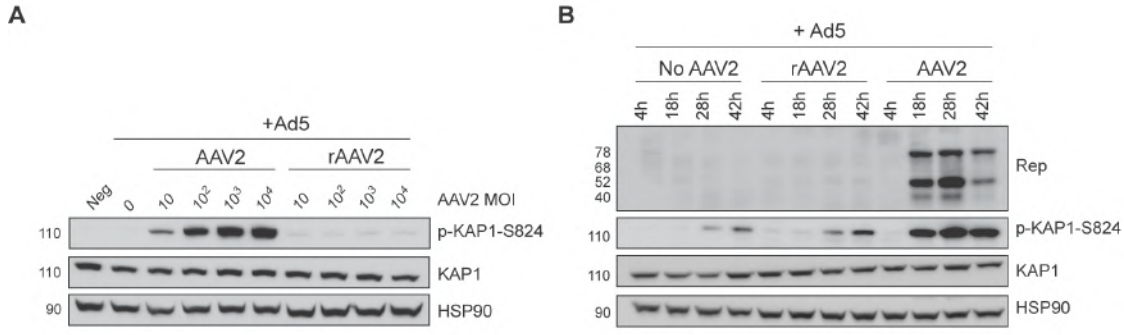


Figure 29. AAV2 replication leads to robust phosphorylation of KAP1-S824
(A) 293T cells infected with increasing MOIs (VG/cell) of either AAV2 or rAAV2 in the presence of Ad5 (2 PFU/cell)) were monitored for levels of p-KAP1-S824 42h after infection by western blotting. **(B)** 293T cells infected with Ad5 alone (2 PFU/cell), or coinfected with Ad5 and either AAV2 or rAAV2 (1000 VG/cell) were monitored for p-KAP1-S824 levels by western blot at 4h, 18h, 24h, and 42h after infection.

5.2.2 The AAV2 Rep proteins Rep52 and Rep78 mediate phosphorylation of KAP1 independently of ATM activation

The observed Rep-KAP1 interaction and correlation between the onset of Rep expression and the phosphorylation of KAP1-S824 led us to ask whether the Rep proteins might be directly modulating KAP1 activity. To address this, cells transfected with various Rep proteins were monitored for p-KAP1-S824 levels by western blotting 27h after transfection. The large Rep proteins possess endonuclease activity shown to trigger DNA damage, and all four Rep proteins share a helicase domain with the potential to also disrupt DNA. In order to minimize the possibility of DDR-dependent induction of p-KAP1-S824, endonuclease mutants (Y156F) of Rep68 and Rep78, and catalytic ATPase mutants (K340H) of all four Rep proteins were used (Chejanovsky and Carter, 1990; Dutheil et al., 2014). Phosphorylation of KAP1-S824 was evident in the presence of both Rep52 and Rep78, independently of either endonuclease or ATPase activity (Figure 30A). Basal levels of p-KAP1-S824 were also visible with Rep40 and Rep68, but were 3- to 6-fold lower than for Rep52 and Rep78. These data suggest that the Rep proteins, in particular Rep52 and Rep78, actively mediate phosphorylation of KAP1-S824 via an unknown, DDR-independent mechanism.

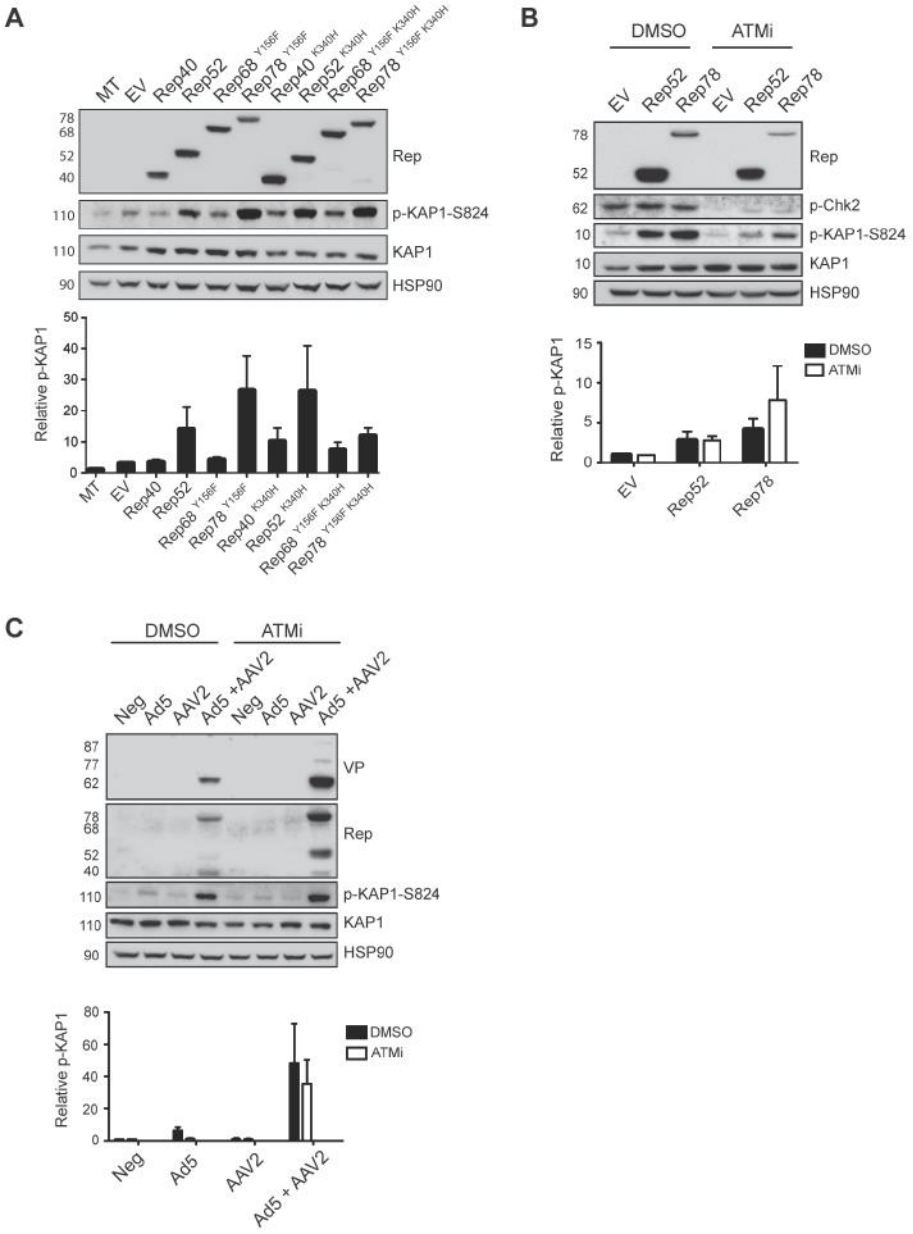


Figure 30. Rep 52 and Rep78 mediate phosphorylation of KAP1-S824 independently of ATM activation.

(A) 293T cells transfected with the various Rep proteins were blotted for p-KAP1-S824 levels by western blotting 27h after transfection. Rep ATPase mutants (K340H) and endonuclease mutants (Y156F) were used to minimize the possibility of DDR-mediated KAP1 phosphorylation. **(B)** 293T cells pretreated with either DMSO or an ATM inhibitor (ATMi) were transfected with and empty vector (EV), Rep52, or Rep78. Cells were harvested 48 h after transfection for analysis of p-KAP1-S824 levels by western blotting. Levels of phosphorylated Chk2 (p-Chk2) were monitored to assess efficiency of ATM inhibition. p-KAP1-S824 levels were quantified using ImageJ software and were normalized to Rep levels to correct for differences in transfection efficiency as a result of pretreatment with ATMi. Values are reported as mean±SEM, n=4. **(C)** DMSO or ATMi treated 293T cells were infected with Ad5 (2 PFU/cell) alone, AAV2 (10 IU/cell) alone, or coinfected with Ad5 and AAV2. Cells were harvested 48h after infection for analysis of p-KAP1-S824 levels by western blotting. p-KAP1-S824 levels were quantified using ImageJ software. Values are reported as mean±SEM, n=4.

To further confirm the independence of this response from the conventional ATM-mediated pathway, we repeated these experiments in the presence of an ATM inhibitor (ATMi). 293T cells were treated with either DMSO or ATMi and then transfected 4 hours later with Rep52 or Rep78, and p-KAP1-S824 levels were monitored 27h after transfection by western blotting. Phosphorylated Chk2 (p-Chk2) was also monitored to control for the efficiency of ATM inhibition. Due to difficulties achieving similar expression levels of Rep in DMSO- and ATMi-treated cells (Figure 30B, lanes 2 and 5, lanes 3 and 6), levels of p-KAP1-S824 were normalized to Rep expression rather than HSP90 for the purpose of quantification. After normalization, no significant difference in Rep-mediated p-KAP1-S824 levels was apparent in ATMi treated cells as compared to controls (Figure 30B). We also performed this experiment in the context of AAV2 infection. 293T cells pretreated with either DMSO or ATMi were infected with Ad5, AAV2, or Ad5 and AAV2 and monitored for p-KAP1-S824 levels by western blotting 27h after infection. Here as well, no difference in p-KAP1-S824 levels were apparent in coinfected cells pretreated with ATMi as compared to controls (Figure 30C). Taken together, these data suggest that increased levels of phosphorylated KAP1-S824 during productive AAV2 infection is supported by the AAV2 Rep52 and Rep78 proteins independently of ATM activation.

Upon induction of DNA double strand breaks (DSB), ATM-mediated p-KAP1-S824 rapidly localizes to sites of damage, forming foci with other mediators of repair such as 53-BP1 and γ-H2AX (Noon, 2010; Goodarzi, 2008). To gain more insight into the possible mechanism through which the Rep proteins mediate phosphorylation of KAP1-S824, we used confocal microscopy to visualize the intranuclear localization of Rep78 and p-KAP1-S824. 293T cells pretreated with ATMi or DMSO were transfected with Rep78-GFP, or an empty vector (EV) control. Cells were also infected with Ad5 in order to trigger the formation of replication centers (RC), where we hoped to identify colocalization of Rep78-GFP and p-KAP1-S824. Robust pan-nuclear phosphorylation of KAP1-S824 was apparent in the presence of Rep78-GFP, both in cells treated with ATMi and controls, again confirming our previous observation that ATM activation is not essential for Rep-mediated p-KAP1-S824 (Figure 31). No difference in intranuclear localization

was observed for Rep78-GFP in cells infected with or without Ad5 however, indicating that Ad5 infection failed to trigger the formation of RCs by Rep78-GFP. Equally, no difference was observed in p-KAP1-S824 localization in the presence of Ad5 infection. Without triggering the formation of RCs, it is impossible to judge the degree of colocalization between Rep and p-KAP1-S824, as both display a pan-nuclear pattern of staining. It will be necessary to repeat these experiments in the context of wild type AAV2 infection as it is likely that the presence of actively replicating AAV2 genomes is necessary both to trigger the formation of RCs as well as to provide a site on which foci of p-KAP1-S824 are able to form. Various time points will also need to be tested. Localization of p-KAP1-S824 to DSB foci has been shown to occur within minutes of assault and to dissipate rapidly into pan-nuclear staining, as observed here. Interestingly, low levels of p-KAP1-S824 were also noticeable in cells infected with Ad5 alone, which were lost in cells treated with an ATM inhibitor, lending credence to the notion that the Rep proteins and Ad5 lead to the phosphorylation of KAP1-S824 via two independent pathways.

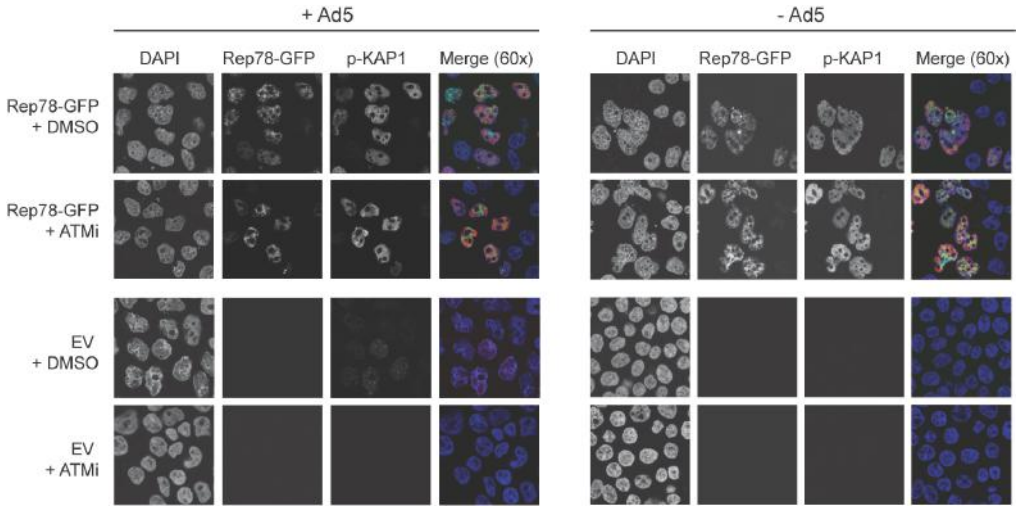


Figure 31. Immunofluorescence analysis of Rep78-GFP and p-KAP1-S824 intracellular localization.
DMSO or ATMi-treated 293T cells were transfected with Rep78-GFP or an empty vector control (EV), with or without Ad5 (2 PFU/cell) infection (left and right, respectively). Cells were fixed 24h after Ad5 infection for immunofluorescence analysis of Rep78-GFP and p-KAP1-S824 intracellular localization.

5.2.3 Phosphorylation of KAP1 by Rep52 and Rep78 is dependent on their shared zinc finger domain.

Given that Rep52 and Rep78 share a C-term zinc finger (ZNF) domain not present in Rep40 or Rep68, we suspected this region might be important for the phosphorylation of KAP1-S824. As the Rep proteins have no kinase activity, it is probable that they mediate phosphorylation of KAP1-S824 through the actions of an intermediary protein, through the coopting of a kinase or inhibition of a phosphatase for example. Indeed, the Rep ZNF domain has already been implicated in several protein interactions, including the recruitment of the kinase PKA. With this in mind, we generated a series of C-terminal truncation mutants in which the ZNF domain was progressively removed (Figure 32A). These were expressed in 293T cells, and p-KAP1-S824 levels were monitored by western blotting 27h after transfection. Strikingly, even the smallest truncation (Δ43) completely abrogated phosphorylation of KAP1-S824. This was true of both Rep52 and Rep78 (Figure 32B). As this effect was so profound, we next questioned whether the Rep proteins might in fact be interacting directly with KAP1 through the ZNF domain. We performed CL-IP for FLAG-tagged proteins in 293T cells coexpressing each of the Rep52 truncation mutants with either FLAG-GFP or FLAG-KAP1. All three of the truncations mutants readily interacted with FLAG-KAP1 but not with FLAG-GFP (Figure 32C), confirming that the Rep proteins are acting through an intermediary protein to mediate the observed phosphorylation of KAP1-S824.

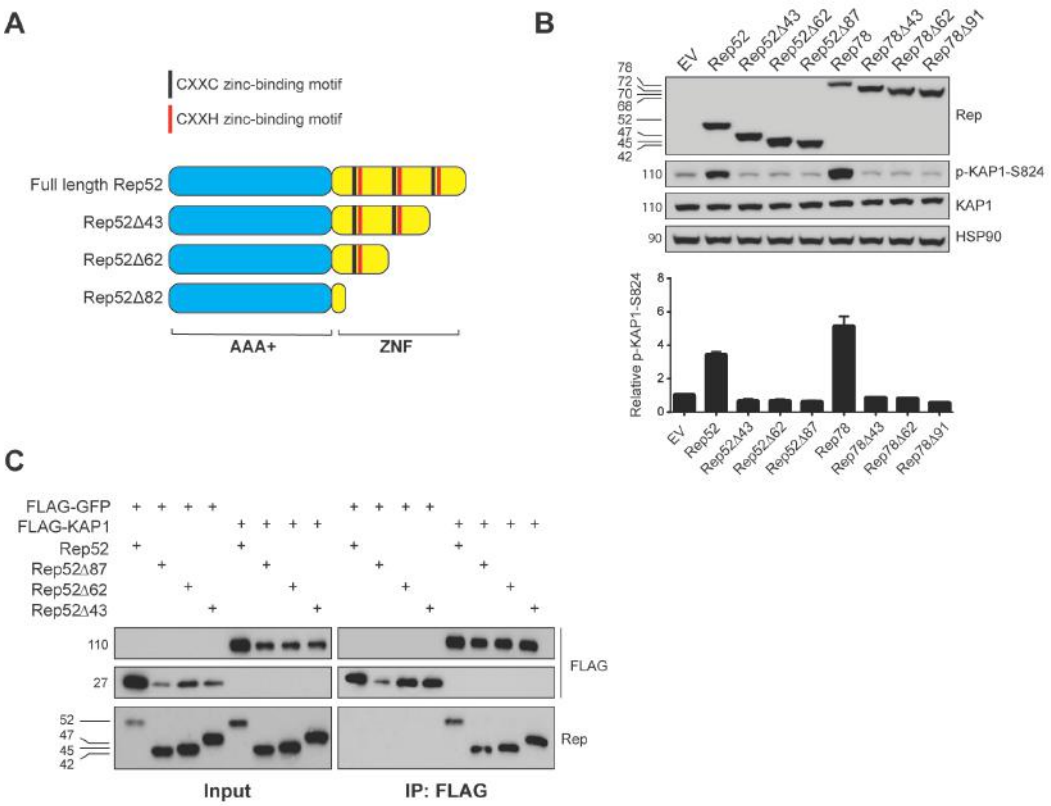


Figure 32. Phosphorylation of KAP1-S824 by Rep52 and Rep78 is dependent on the shared ZNF domain.
(A) Schematic diagram representing full length Rep52, comprising an ATPase domain (AAA+) shown in blue and zinc finger domain (ZNF) shown in yellow, and the C-terminal truncation mutants in which the ZNF domain is progressively removed. Black bars indicate a CXXC zinc-binding motif, and red bars a CXXH zinc-binding motif. **(B)** 293T cells transfected with Rep52, Rep78, various ZNF truncation mutants or an empty vector control (EV) were harvested for analysis of p-KAP1-S824 levels 48h after transfection by western blotting. **(C)** Cross-linked IP for FLAG-tagged proteins in lysates from 293T cells transfected with FLAG-KAP1, or a FLAG-GFP control, and Rep52 or the Rep52 ZNF truncation mutants. Values are reported as mean±SEM, n=3.

5.2.4 Rep52 and Rep78 interact with the phosphatase PP1 and its negative regulator NIPP1 to enhance levels of phosphorylated KAP1

Potential cellular factors that could be interacting with Rep to mediate phosphorylation of KAP1-S824 include protein phosphatase 1 (PP1) and its specific inhibitors. Upon completion of DNA repair, basal levels of p-KAP1-S824 are restored through the combined activities of PP1α, PP1β, and protein phosphatase 4 (Li et al., 2010; Pfeifer, 2012). Several regulatory subunits of PP1 were identified as interaction partners for Rep alongside KAP1 in our original BioID screen, including the nuclear inhibitor of PP1

(NIPP1/PPP1R8), which led us to hypothesize that the Rep proteins could be interfering with this pathway. We postulated that Rep was inhibiting PP1 activity by bridging the interaction between PP1 and its negative regulator NIPP1, thus leading to increased nuclear levels of p-KAP1-S824.

To test this hypothesis, we first searched for potential PP1 binding sites within the Rep sequence. Using the conserved consensus binding sequence [KR][X_{0-1}][VI]{P}[FW] as our guideline (Meiselbach et al., 2006), we discovered one putative binding site in the ATPase domain (K372-W376), partially overlapping with the Walker B motif (Figure 33A). Co-IP experiments in cells expressing FLAG-PP1α and Rep52 confirmed the physical interaction between Rep and PP1α (Figure 33B). However, as both PP1-α and –β share the same substrate binding domain, it is impossible to know from this experiment which of the isoforms Rep might interact with in a physiological context. In order to assess the relevance of this interaction with respect to the Rep-mediated phosphorylation of KAP1-S824, we introduced a lysine to alanine point mutation at the first lysine in the putative PP1 binding site (K372A). This was the only possible site in which to introduce a mutation due to the overlap between the consensus PP1 binding site and Walker B motif. Surprisingly, the PP1-binding mutant Rep52^{K372A} also physically interacted with FLAG-PP1α (Figure 33C). Interestingly however, when co-IP experiments were repeated with the inclusion of the Rep ATPase mutant and PP1-binding/ATPase double mutant (Rep52^{K340H} and Rep52$^{K340H\ K372A}$), Rep52^{K340H} again interacted with FLAG-PP1α, while the interaction between FLAG-PP1α and Rep52$^{K340H\ K372A}$ was lost (Figure 33C). Importantly, mutation of the PP1 binding site alone was sufficient to completely abrogate Rep-mediated phosphorylation of KAP1-S824 (Figure 33D).

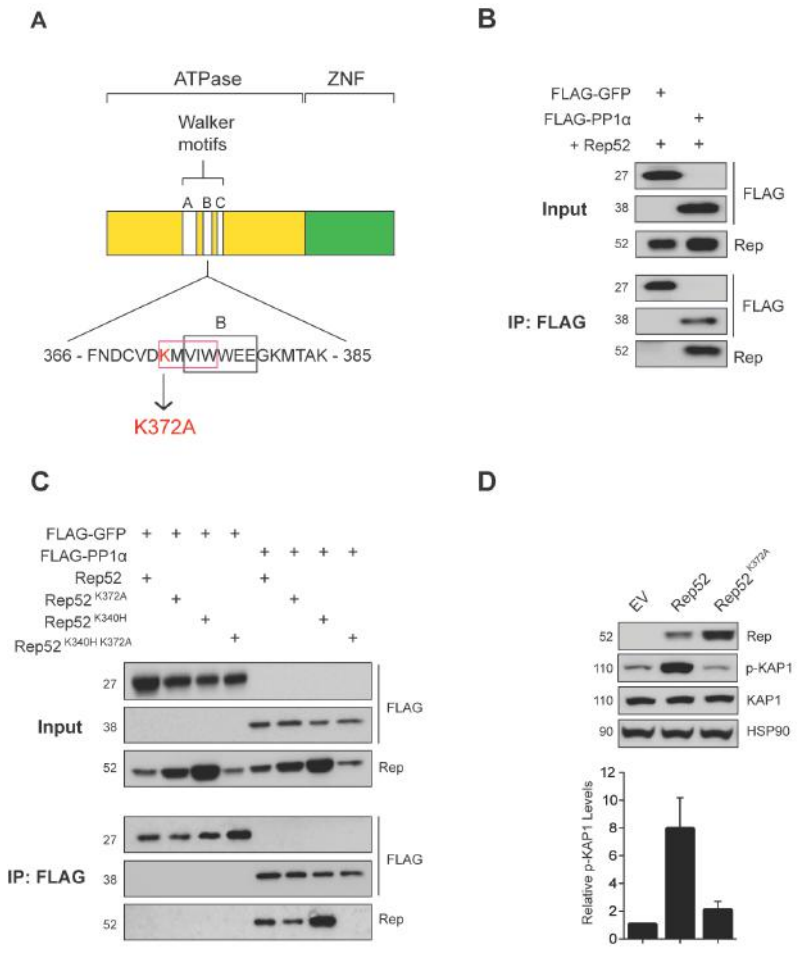


Figure 33. Rep52 binds the protein phosphatase PP1α, and mutation of the binding site abrogates Rep-mediated phosphorylation of KAP1-S824.
(A) Schematic diagram depicting the putative PP1-binding site in the Rep ATPase domain. The Walker B motif is outlined in black, and the partially overlapping consensus binding site is outlined in pink. Lysine 372 in position 1 of the binding site is shown in red. **(B)** Co-IP for FLAG-tagged proteins in lysates from 293T cells transfected with Rep52 and either FLAG-PP1α or FLAG-GFP control. **(C)** Co-IP for FLAG-tagged proteins in lysates from 293T cells transfected with either FLAG-PP1α or FLAG-GFP control, and Rep52, the PP1 binding mutant Rep52^{K372A}, the ATPase mutant Rep52^{K340H}, or the double mutant Rep52$^{K340H\ K372A}$. **(D)** 293T cells transfected with empty vector control (EV), Rep52, or Rep52^{K372A} were analyzed 27h after transfection for p-KAP1-S824 levels by western blotting.
Values are reported as mean±SEM, n=3.

Given the proximity of the K372A mutation to the Walker B motif, we deemed it necessary to ensure that the ATPase activity of Rep52^{K372A} had not been compromised. Transcriptional regulation by the Rep proteins is dependent on a functional ATPase domain, and so we addressed this

concern by assessing the ability of Rep52^{K372A} to regulate expression from the AAV2 p5 promoter. The previously described p5-mCherry reporter construct was cotransfected with either wild-type Rep52 or Rep52^{K372A}, and p5 activity was determined 48 h after transfection by western blotting for mCherry. Repression of p5 by the PP1 binding mutant Rep52^{K372A} was equal to that of wild type Rep52 (Figure 34A), confirming that the loss of KAP1-S824 phosphorylation by Rep52^{K372A} was not an off-target effect of having disrupted ATPase function. In addition, we were able to demonstrate by CL-IP that the K372A mutation did not simply interfere with the interaction between Rep and KAP1 (Figure 34B), further supporting the hypothesis that the loss of a functional interaction between Rep and an intermediary protein, potentially PP1, is responsible for the reduced phosphorylation of KAP1-S824 in the presence of Rep52^{K372A}.

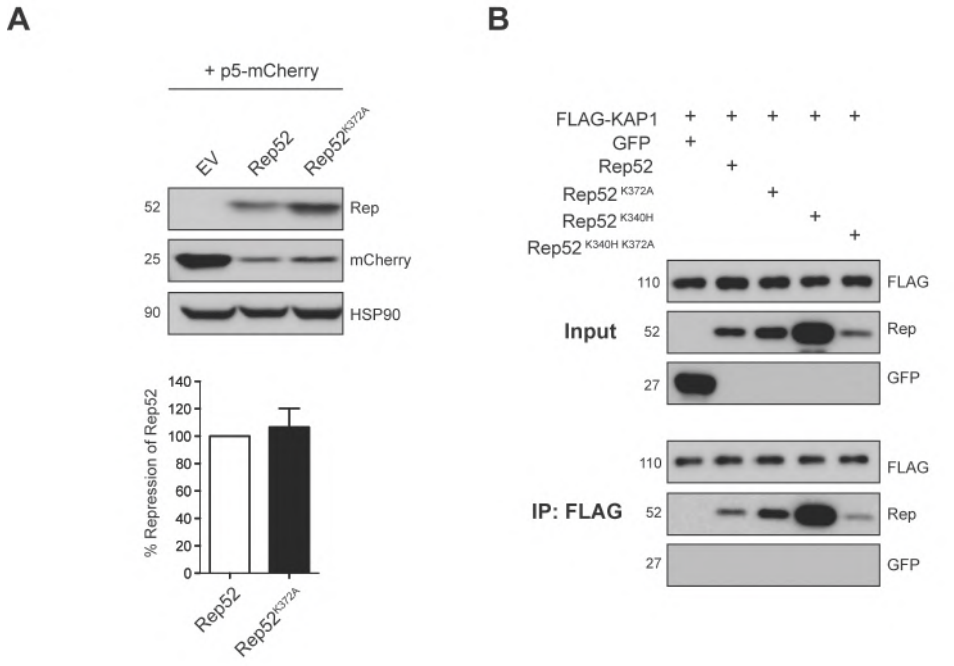


Figure 34. Validation of the Rep-K372A PP1-binding mutant.

(A) 293T cells transfected with a p5-mCherry reporter construct and Rep52, Rep52^{K372A}, or empty vector control (EV) were harvested 48h after transfection. p5 activity was determined by western blotting for mCherry. Values are reported as mean±SEM, n=3 **(B)** Cross-linked IP for FLAG-tagged proteins from lysates of 293T cells transfected with FLAG-KAP1 and GFP control, Rep52, PP1 binding mutant Rep52^{K372A}, ATPase mutant Rep52^{K340H}, or the double mutant Rep52$^{K340H\ K372A}$.

Based on our earlier observation that the Rep ZNF domain is also crucial for the phosphorylation of KAP1-S824, we next asked whether this domain could be recruiting NIPP1 and forming a complex with PP1 bound at the ATPase domain. To address this, we performed CL-IP with lysates from cells expressing NIPP1 with Rep52, each of the Rep52 ZNF truncation mutants, or the PP1-binding mutant Rep52^{K372A}. NIPP1 readily interacted with both Rep52 and Rep52Δ42, while a progressive loss of interaction was apparent with the two larger truncations, Rep52Δ63 and Rep52Δ87 (Figure 35). This displayed a degree of variation however, ranging from a complete loss of binding to only a modest decrease as compared to Rep52Δ42. Interestingly, the interaction with NIPP1 was consistently lost with Rep52^{K372A}. Taken together, these data suggest that PP1 may in fact represent the main binding partner for Rep, and that NIPP1 is recruited as a consequence. It is conceivable that the ZNF domain acts to stabilize the PP1-NIPP1 holoenzyme, leading to prolonged inhibition of PP1 and increased levels of nuclear p-KAP1-S824.

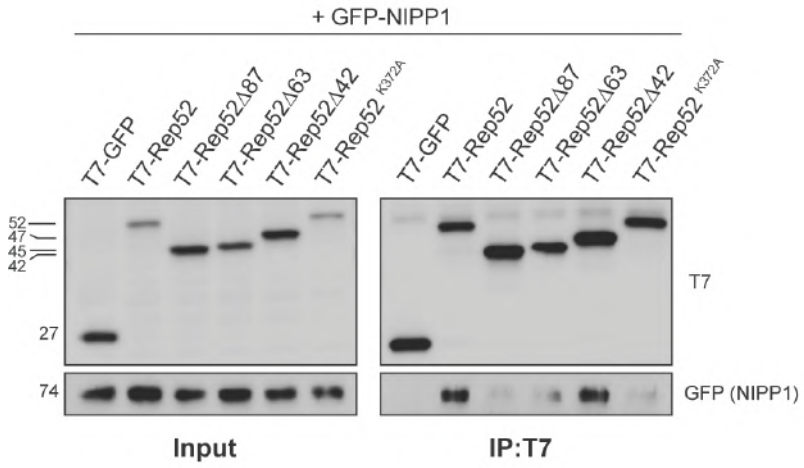


Figure 35. Rep interaction with the PP1 inhibitor NIPP1 is dependent on PP1 binding and intact Rep ZNF domain.
Cross-linked IP for T7-tagged proteins in lysates from 293T cells transfected with GFP-NIPP1, and T7-GFP, T7-Rep52, T7-Rep52 ZNF truncation mutants, or the PP1 binding mutant T7-Rep52^{K372A}.

5.2.5 Inhibition of PP1 is essential for AAV2 transcription and replication

To determine the functional relevance of the Rep-PP1-NIPP1 interaction in the AAV2 life cycle, we performed replication experiments with an AAV2 plasmid containing the PP1 binding mutation (Rep-K372A). The wild-type plasmid, pAV2 (Rep-WT), was used as a positive control, and the NTP-binding/ATPase mutant K340H (Rep-K340H), which does not support AAV2 replication (Chejanovsky and Carter, 1990), was used as a negative control. Cells were transiently transfected with each of the infectious plasmids and infected with Ad5 4h after transfection. Cells were then harvested 72h after transfection for quantification of vector genomes, and *rep* and *cap* transcription levels. As expected, Rep-WT fully supported AAV2 genome replication in the presence of Ad5, while no response was apparent in the presence of either Rep-K340H or Rep-K372A (Figure 36A). For the transcriptional analysis, primer sets targeting p5, p5 + p19, and p5 + p19 + p40 transcripts were used – representing respectively the large Rep proteins alone, all Rep proteins, and Rep and Cap proteins together – as the use of a single polyadenylation signal by AAV RNAs precludes the analysis of p19 and p40 transcripts separately from p5 transcripts. We observed an increase in mRNA levels in response to Ad5 coinfection with all three primer sets in the presence of Rep-WT (Figure 36B). Base-line transcription levels were substantially higher with Rep-K340H than with Rep-WT due to the reduced capacity for transcriptional regulation of this Rep mutant (Kyostio et al., 1995), however no response to Ad5 infection was apparent. Similarly, no response to Ad5 infection was observed in the presence of Rep-K372A (Figure 36B). Interestingly, base-line expression levels from p19 and p40, but not p5, were elevated for Rep-K372A, albeit at significantly lower levels than for Rep-K340H.

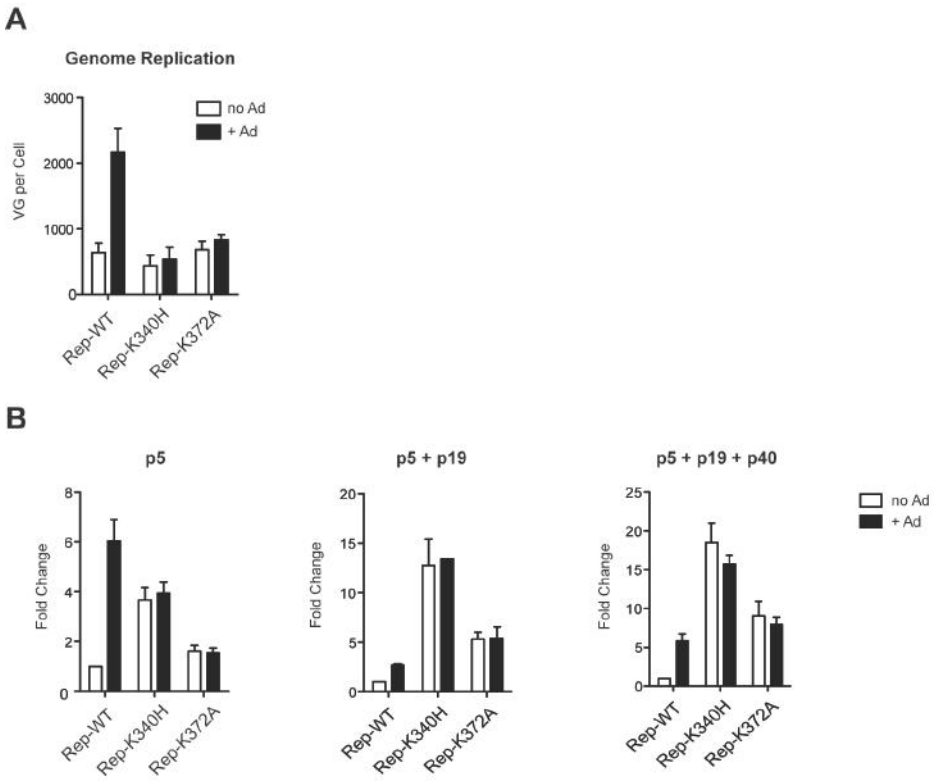


Figure 36. Interaction between Rep, PP1, and NIPP1 is essential for AAV2 replication and transcription.
(A-B) 293T cells transfected with pAV2-WT (Rep-WT), pAV2-K340H (Rep-K340H), or pAV2-K372A (Rep-K372A), and infected with Ad5 (4 PFU/cell) 4h after transfection were harvested 72h after transfection for DNA and RNA extraction. **(A)** Real time qPCR analysis of AAV2 replication. **(B)** RT-qPCR analysis of AAV2 gene transcription under the same conditions. p5 (left panel) indicates that the data is specific for p5 transcripts; primers that bind after p19 but before p40 (middle panel) detect transcripts generated from both p5 and p19; p40 primers (right panel) detect transcripts from all three promoters. Fold change was calculated relative to the transcript levels in the presence of WT Rep but in the absence of Ad5.
Values are reported as mean±SEM, n=3.

To test this further, we next performed replication experiments in 293T cells depleted for KAP1 to determine whether replication in the presence of Rep-K372A could be rescued. Control (shEMPTY) and KAP1-depleted (shKAP1) 293T cells were transfected with either Rep-WT or Rep-K372A and infected with Ad5 as described above. Cells were then harvested 72h after transfection for quantification of vector genomes and *rep* and *cap* transcription levels. As shown in the first set of experiments, Rep-WT fully supported replication in the presence of Ad5, and while the effect was less pronounced under these conditions than in our original replication experiments, Rep-WT-supported replication was enhanced ~2-fold in KAP1-depleted cells as compared to controls (Figure 37A). Interestingly, a very

modest increase in replication was also detectable in KAP1-depleted cells transfected with Rep-WT in the absence of Ad5 infection as compared to controls, although these experiments will need to be repeated to determine the significance of this observation. As expected, Rep-K372A was unresponsive to Ad5 infection in control cells. Surprisingly this was also true in KAP1-depleted cells, indicating that depletion of KAP1 is not sufficient to rescue replication by Rep-K372A (Figure 37A). This was confirmed at the transcriptional level, where Rep-K372A was unresponsive to Ad5 infection in control as well as KAP1-depleted cells with all primer sets used (Figure 37B). Modest disregulation of p19 and p40 was again observed in the presence of Rep-K372A. Transcription in the presence of Rep-WT however did not correlate with the genome replication data; p5 transcription in response to Ad5 infection appeared reduced in KAP1-depleted cells as compared to controls, and no significant difference between controls and KAP1-depleted cells was detectable for p19 and p40 transcription.

Taken together, these data suggest that, while the interaction of Rep with PP1 and the associated enhancement of p-KAP1-S824 levels appear to play a role in AAV2 replication, the mechanism is more complex than we yet understand. It is possible that the minimal levels of KAP1 still present in KAP1-depleted cells are sufficient to maintain repression of the AAV2 genome in the context of Rep-K372A. It is also possible that phosphorylated KAP1-S824 has an entirely separate function of which we are not yet aware. Alternatively, the observed disregulation of p19 and p40 in the presence of Rep-K372A suggests that additional interactions with unknown factors involved in the transcriptional regulation of these promoters may be interrupted.

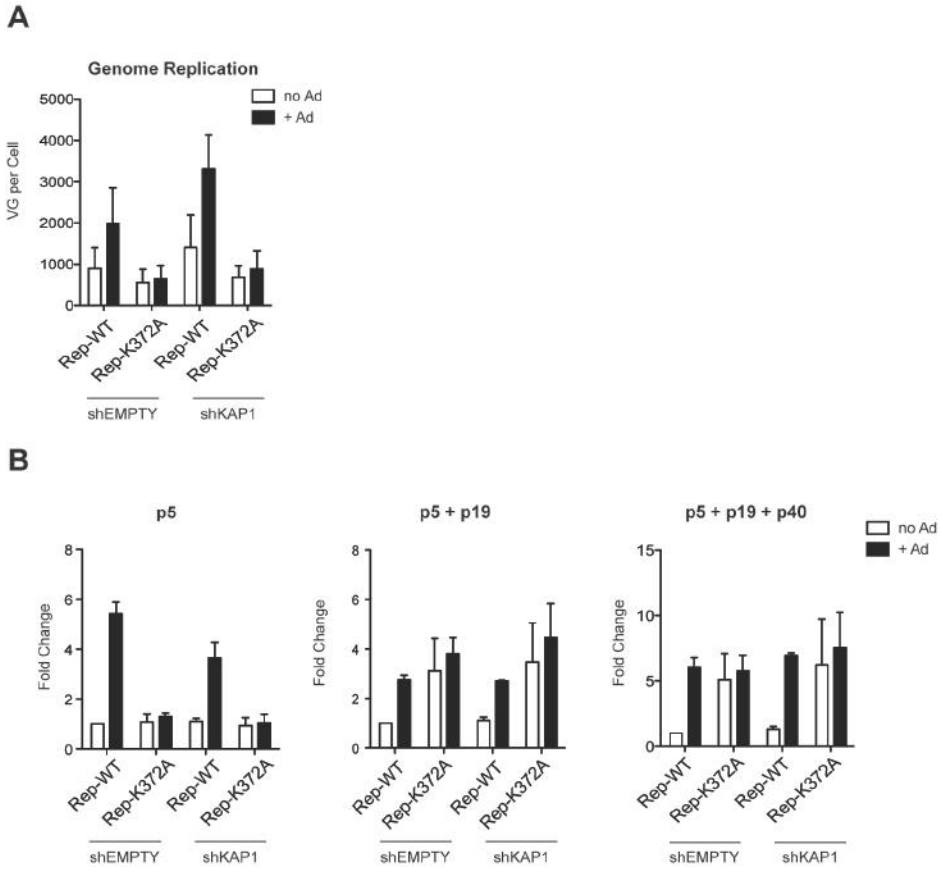


Figure 37. KAP1 depletion does not rescue Rep-K372A-supported replication
(A-B) Control and KAP1-depleted 293T cells transfected with pAV2-WT (Rep-WT) or pAV2-K372A (Rep-K372A) and infected with Ad5 (4 PFU/cell) 4h after transfection were harvested 72h after transfection as described above. **(A)** qPCR analysis of AAV2 replication. **(B)** RT-qPCR analysis of AAV gene expression as described for Figure 36.
Values are reported as mean±SEM, n=3.

5.3 Discussion

In the context of DNA damage, the ATM kinase phosphorylates KAP1-S824, resulting in dispersion of the repressive complex, relaxation of heterochromatin, and relief from transcriptional repression, after which basal levels of p-KAP1-S824 are restored by the phosphatases PP1α/β and PP4 (Cheng et al., 2014). Reactivation of both CMV and KSHV is characterized by the phosphorylation of KAP1-S824, mediated respectively by the cellular kinase mTOR and the viral kinase vPK (Chang et al., 2009; Rauwel et al., 2015). Similarly, we show here that AAV2 lytic replication is associated with a significant increase in p-KAP1-S824 levels, suggesting that KAP1 corepressor activity must be inactivated for replication to take place. In the context of CMV, the forced phosphorylation of KAP1-S824 by the ATM activator chloroquinine results in the reactivation of latent CMV from non-permissive CD34+ HSCs, as demonstrated by increased levels of viral RNA and DNA, the presence of IE antigens, and the release of replication competent virus into the supernatants (Rauwel et al., 2015). The levels of gene expression and DNA accumulation in this context are still significantly lower than what is observed when CMV-harboring precursors are differentiated into permissive macrophages or DCs however. The investigators showed that treatment of cells with tumor necrosis factor α (TNFα) further boosted viral RNA levels and virion production by a factor of 10 as a result of NF-κB induction, suggesting that, although necessary to remove repressive marks from viral chromatin, phosphorylation of KAP1-S824 is not sufficient to fully reactivate the viral genome. It is likely that AAV2 is subject to a similar scenario, where phosphorylation of KAP1-S824 is necessary to render the viral genome permissive to transcription by removing repressive H3K9me3, but that transcription is only initiated in the presence of helper virus factors. Indeed, this model is supported by our earlier observation that KAP1 depletion, which is equivalent to p-KAP1-S824 with respect to heterochromatinization, was not sufficient to initiate any observable AAV2 transcription or replication in the absence of Ad5.

Another interesting observation made during the forced reactivation of CMV through ATM activation was that elevated levels of p-KAP1-S824 were only visible in cells that were also positive for CMV IE antigens (Rauwel et

al., 2015). The investigators interpreted this to mean that KAP1-S824 was phosphorylated through the combined activities of ATM and some CMV-encoded factor. Similarly, the direct phosphorylation of KAP1-S824 by the KSHV kinase vPK was shown to enhance viral reactivation (Chang et al., 2009). Interestingly, further analysis of vPK substrates revealed overlapping specificity between vPK and ATM kinases, suggesting that the relevance of vPK for viral reactivation may be dependent on the strength of ATM activation (Chang et al., 2009). Here, we also show that AAV2 lytic replication is associated with significantly enhanced levels of p-KAP1-S824. Furthermore, we show that this phosphorylation is correlated with interactions between Rep, the phosphatase PP1, and its negative regulator NIPP1, suggesting that Rep may act as a PP1 inhibitor to increase nuclear levels of p-KAP1-S824. Like KSHV vPK, the large Rep proteins have been shown to induce low levels of phosphorylation of the DDR factors ATM, H2AX, and SMC1 (Schwartz et al., 2009). It is interesting to consider what role this might have during AAV2 and Ad5 coinfection, where Ad5-mediated degradation of the MRN complex and the resulting inactivation of the DDR NHEJ pathway is necessary for both Ad5 and AAV2 replication (Schwartz et al., 2007; Stracker et al., 2002). It is conceivable that, under these conditions, the Rep proteins play an important role in phosphorylating certain downstream substrates of MRN activation that may be beneficial for AAV replication, such as KAP1. Indeed, while ATM activation is detrimental to rAAV replication, it has been shown to be necessary for efficient wtAAV replication (Collaco et al., 2009; Schwartz et al., 2007), supporting a role for a downstream effector in the efficient propagation of wtAAV.

It is also noteworthy that KSHV, CMV, and AAV2 – three biphasic DNA viruses – are all targets of KAP1 while also seeming to contain countermeasures to KAP1 repression. This suggests that these viruses may in fact have domesticated KAP1 repression in order to preserve their latency until conditions for replication are met. Like all herpesviruses, KSHV and CMV establish latent infections, persisting as nuclear episomal DNA. Although still a matter of some debate, AAV latency can be established through integration into the host genome, or also through long-term

episomal persistence. Mounting evidence suggests that innate sensing of foreign DNA occurs not only in subcellular compartments in which DNA should not exist, but also within the nucleus itself. The interferon-inducible protein IFI16 is an innate DNA sensor (Unterholzner et al., 2010) that localizes predominantly to the nucleus in non-immune cells (Li et al., 2012; Orzalli et al., 2012), where it has been shown to induce responses to replicating CMV and KSHV (Kerur et al., 2011; Li et al., 2013). Interestingly, IFI16 has also been shown to mount responses in B cells containing latent episomal forms of KSHV or EBV (Ansari et al., 2013; Singh et al., 2013). IFI16 recognizes foreign DNA by tracking along stretches of exposed dsDNA to assemble supramolecular signaling platforms (Stratmann et al., 2015), and nucleosome formation presents a barrier to these signaling complexes. Viral nucleosomes are often generated with slightly different conformations than normal host chromatin, (Lieberman, 2008), resulting in longer stretches of unprotected DNA and creating the basis for the distinction between self and non-self (Stratmann et al., 2015). It is therefore possible that heterochromatinization by KAP1 allows latent episomal AAV to better evade immune recognition by reducing stretches of exposed viral DNA. Heterochromatinization may also help to prevent deleterious recombination events, both intramolecularly between the ITRs, and intermolecularly with other viral episomes or the host genome, as well as to protect the latent viral genome from degradation.

Upon the induction of lytic replication through helper virus coinfection, inhibition of PP1 by Rep52 and Rep78 leads to enhanced levels of p-KAP1-S824. It is not yet entirely clear how Rep achieves this inhibition. We clearly show that Rep52 and PP1 physically interact, however it was necessary to mutate the putative PP1-binding site in the background of the additional NTP-binding mutation (K340H) in order to abolish this interaction. There is precedence to suggest that a functional interaction between PP1 and its regulatory subunits is based on multiple points of interaction, only one of which necessarily consists of the conserved binding site. It has also been demonstrated that mutation of only one interaction site may be sufficient to abolish the regulation of PP1 while being insufficient to abolish the actual

physical association between PP1 and its regulatory component (Beullens et al., 2000; O'Connell et al., 2012). Our results suggest the existence of a second interaction site, either defined by the Rep NTP-binding site itself or created by the conformational change/dimerization associated with NTP-binding, which is sufficient to mediate a physical interaction between Rep and PP1 but not sufficient to inhibit PP1 activity. It is possible that direct binding of Rep to PP1 simply obstructs the substrate binding site necessary for interaction with KAP1, however the apparent role for both PP1 and the Rep ZNF domain in the interaction between Rep and NIPP1 suggests that the three proteins act as a complex, with the ZNF potentially stabilizing the PP1-NIPP1 holoenzyme and prolonging PP1 inhibition.

To complicate matters further, although first identified as a potent inhibitor of PP1 (Beullens et al., 1992), NIPP1 does not systematically inhibit PP1 activity. NIPP1 is one of the evolutionarily oldest regulators of PP1, and more than 1/3 of the nuclear pool of PP1 forms a holoenzyme with NIPP1 (Jagiello et al., 1995). NIPP1 contains three functional domains: (1) an N-terminal Forkhead Associated (FHA) domain, which specifically binds phosphorylated threonine (p-Thr) residues followed by proline, (2) a central PP1-binding domain containing the consensus PP1-binding motif, and (3) a multifunctional C-terminal domain that binds RNA, has endoribonuclease activity, and inhibits PP1 activity via an unknown mechanism (Beullens et al., 2000; Jagiello et al., 1995, 1997). Depending on (1) highly dynamic electrostatic interactions between NIPP1 and PP1, and (2) the recruitment of substrates by either the NIPP1 FHA domain or by PP1 itself, NIPP1 can either enhance or inhibit PP1 phosphatase activity (O'Connell et al., 2012). It is conceivable then that the NIPP1 FHA domain binds KAP1, effectively recruiting PP1 to mediate KAP1-S824 dephosphorylation, and that Rep binding to PP1 interferes with this activity. This is unlikely however as KAP1 does not contain the necessary p-Thr/Pro residues for binding by the NIPP1 FHA domain, and furthermore we could not detect an interaction between KAP1 and NIPP1.

We show here that the interactions between Rep, PP1, and NIPP1 can mediate the phosphorylation of KAP1-S824 independently from ATM activation, however this does not exclude a role for the activation of effectors

of the cellular DDR in the context of a productive infection. In fact, the nature of the mechanism outlined above presupposes an initial trigger for KAP1 phosphorylation, and productive AAV2 infection is known to trigger the robust activation of both ATM and DNA-PK. (Schwartz et al., 2009), both of which are shown to mediate phosphorylation of KAP1-S824. Taken together, these data suggest a two-part mechanism in which the cellular DDR upon initial viral infection triggers the phosphorylation of KAP1-S824, a signal then potentiated through Rep-mediated antagonism of PP1. This hypothesis may also help to explain the common observation that genotoxic stress of virtually any form enhances AAV2 replication, as this would presumably lead to a DDR-dependent phosphorylation of KAP1-S824.

Finally, we demonstrate that Rep proteins unable to interact with PP1 are also unable to support both viral transcription and DNA replication, providing a potential link between Rep-mediated inhibition of PP1, phosphorylation of KAP1-S824, and AAV2 replication. However, the fact that depletion of KAP1 was not sufficient to rescue replication or induction of transcription in the context of the PP1-binding mutant Rep-K372A suggests that the mechanism involved is more complex than we yet understand. It is possible that low levels of KAP1 that remain after depletion are sufficient to mediate repression of latent AAV2. Alternatively, depletion of KAP1 may not be equivalent to p-KAP1-S824 in the context of AAV2; in other words, p-KAP1-S824 may have an independent function of which we are not currently aware. These observations together with the apparent deregulation of transcription from p19 and p40 in the presence of Rep-K372A support a role for KAP1 in proximal promoter pausing at these promoters. Depletion of KAP1 results in a decrease in magnitude of induction from KAP1-7SK snRNP-regulated primary response genes (PRGs) upon stimulation, suggesting that KAP1 is necessary for the proper induction of PRGs (McNamara et al., 2016). One might imagine that KAP1-7SK snRNP complexes regulate pausing of RNA pol II at p19 and p40, maintaining these promoters poised for rapid activation upon stimulation by helper virus coinfection; this would render p19 and p40 similar to cellular PRGs and thus explain why KAP1 depletion would not enhance stimulation of transcription

from these promoters. Arguing against this possibility however is the fact that no reduction in stimulation from these promoters was observed in the presence of Rep-WT after KAP1 depletion. At the same time, PP1 has been shown to dephosphorylate and thus deactivate the RNA pol II CTD (Kim et al., 2002). Rep proteins unable to inhibit PP1 may as a consequence lead to more constitutively active RNA pol II at sites of proximal promoter pausing, potentially resulting in the observed deregulation of p19 and p40 transcription in the presence of Rep-K372A.

Alternatively, it is possible that the introduction of the K372A mutation has altered another core function of the Rep proteins, which translates to deregulated transcription. Rep-K372A may not undergo the oligomerization necessary for normal AAV gene regulation, however this should be visible at the p5 promoter as well. Although we showed that Rep52-K372A was as efficient as wild type Rep52 in regulating the AAV2 p5 promoter, this is an indirect validation of Rep52-K372A activity. It would be prudent to further confirm that Rep function is maintained by performing helicase and ATPase function assays, as the interruption of either of these activities could have serious implications. Finally, it is also possible that the K372A mutation interferes in the interaction between Rep and a cellular factor involved in an ATPase domain-dependent mechanism of repression by the Rep proteins, akin to what is thought to occur with the NTP-binding K340H mutants (Dutheil et al., 2014; Kyöstiö et al., 1994). This would lead to deregulation of p19 and p40 without necessarily affecting p5, as binding of the RBS at p5 by the Rep OBD mediates repression independently from the ATPase domain.

In conclusion, in this chapter we asked how AAV2 might be released from KAP1-mediated repression so as to allow for replication to occur. We showed that AAV2 lytic replication is associated with a strong phosphorylation of KAP1-S824, suggesting the KAP1 corepressor activity is inactivated during replication. We presented evidence to suggest that this modulation of KAP1 activity is mediated through complex interactions between AAV2 Rep52 (and presumably Rep78 through the shared ATPase and ZNF domains) with the phosphatase PP1 and its regulator NIPP1, which

lead to enhanced levels of nuclear p-KAP1-S824, and importantly that these interactions were necessary for normal AAV2 transcription and replication upon helper virus coinfection. While these data establish an active role for AAV2 in the regulation of its epigenetic landscape and transcriptional competence, there still remains the question of how AAV2 transitions from being repressed by KAP1 to actively counteracting KAP1 via the Rep proteins. Evidently, basal levels of Rep expression during latency are not sufficient to lift KAP1 repression, which implies the need for some biological switch. In the next chapter, we investigate the role of the AAV2 helper viruses in providing this switch.

Chapter 6. KAP1 targeting by the AAV2 helper viruses Ad5 and HSV-1 as a novel helper function for AAV2 replication with the potential for improving recombinant AAV vector production.

6.1 Introduction

In the previous chapters, we demonstrated that latent AAV2 genomes are bound by the corepressor KAP1, leading to the transcriptionally repressive methylation of AAV2-associated histones. We were also able to show that KAP1 corepressor activity is inactivated through phosphorylation of KAP1-S824 during AAV2 lytic replication through interactions between Rep52/Rep78, PP1, and NIPP1. This suggests that KAP1 repression must be inactivated for AAV2 replication to take place and that the AAV2 Rep proteins play an active role in counteracting this repression. The question remains however of how latent AAV2 transitions from being bound and repressed by KAP1 to actively counteracting this repressor, as the basal levels of Rep expression during latency are clearly not sufficient to mediate this effect. This implies the existence of a biological switch necessary to allow for the upregulation of *rep* expression prior to the onset of p-KAP1-S824 accumulation. As AAV2 is a dependovirus, we hypothesized that helper viruses could potentially fulfill this role.

The minimal set of adenovirus helper proteins for AAV replication is comprised of E1A, E1B55K, E2A, E4orf6, and VA RNA, which act together to directly support viral replication as well as alter the cellular milieu such that it becomes conducive to viral replication. E1A directly interacts with the cellar factor YY1 to relieve repression of the AAV p5 promoter (Chang et al., 1989; Lewis et al., 1995; Shi et al., 1991; Weitzman et al., 1996), the E2A gene product is a ssDNA binding protein necessary for the efficient replication of ssDNA AAV genomes (Ward et al., 1998), and VA RNA binds the kinase PKR to overcome the shutdown of protein translation that is often associated with viral infection (Nayak and Pintel, 2007). Of particular relevance to this study, E1B55K and E4orf6 act in complex as cullin-based E3 ubiquitin ligases to facilitate infection by inducing the degradation of cellular proteins that adversely affect viral replication (Harada et al., 2002; Querido et al.,

2001). E4orf6 forms the core complex with elongins B and C, either Cul5 or Cul2, and Rbx1, after which E1B55K binds to the complex and recruits substrates for ubiquitination (Blanchette et al., 2004; Cheng et al., 2013; Querido et al., 2001). The particular cellular proteins targeted are serotype-specific (Forrester et al., 2011); in the context of Ad5, a Cul5-based complex leads to the degradation of Mre11, DNA ligase IV, Bloom helicase (BLM), p53, MDM2, and Tip60 (Baker et al., 2007; Gupta et al., 2013; Orazio et al., 2011; Querido et al., 2001; Stracker et al., 2002). Degradation of Mre11, DNA ligase IV, and BLM serves to inactive the cellular DDR and NHEJ pathway, while degradation of p53 and MDM2 prevents apoptosis. Tip60 is a lysine acetyltransferase shown to repress transcription from the adenovirus E1A promoter (Gupta et al., 2013). Based on this, we asked whether Ad5 infection might also lead to the degradation of KAP1 via this pathway.

Another adenovirus protein that plays an essential role during infection is E4orf3. This protein associates with the cellular promyelocytic leukemia (PML) protein, the key organizer of PML nuclear bodies (NB) (Carvalho et al., 1995). PML NBs are functionally complex nuclear domains that recruit an astonishing variety of proteins involved in central cellular processes such as DNA replication, transcription, epigenetic silencing, and host defense mechanisms against viral infection (Lallemand-Breitenbach and de Thé, 2010; Mao et al., 2011). Through its interaction with PML, E4orf3 reorganizes PML NBs into distinctive track-like structures where it sequesters numerous cellular proteins involved in DDR and repair pathways (Carson et al., 2009; Doucas et al., 1996; Evans and Hearing, 2005). E4orf3 also binds and reorganizes to these PML tracks the TIF1 family members TIF1α and TIF1γ, which are closely related to KAP1 (TIF1β) (Vink et al., 2012; Yondola and Hearing, 2007). In addition, E4orf3 leads to the degradation of TIF1γ independently of E1B55K or E4orf6 (Forrester et al., 2012). Interestingly, E1B55K was shown to bind all three TIF1 family members, however no reorganization or degradation of KAP1 has previously been observed (Forrester et al., 2012; Vink et al., 2012; Yondola and Hearing, 2007).

Recently however, binding of E1B55K to KAP1 has been suggested to play a role during the early stages of adenovirus infection (Bürck et al., 2015). Adenovirus replication can be restricted at early stages of infection by factors that affect the initial decondensation of the viral genome upon entry, such as the histone deacetylatse Daxx, which is also targeted for degradation by the E1B55K/E4orf6 ubiquitin ligase complex (Schreiner et al., 2010). The protein SPOC-1 (survival time-associated PHD protein in ovarian cancer) is another cellular factor involved in DNA repair that acts via the selective modulation of and cooperation with transcriptionally repressive chromatin modifiers, and which has been suggested to inhibit adenovirus decondensation through an association with KAP1 (Bürck et al., 2015; Schreiner et al., 2013a). Significantly reduced levels of E1B and E4orf6 were apparent in adenovirus-infected cells overexpressing KAP1, suggesting KAP1 can inhibit adenovirus gene expression, however depletion of KAP1 had no effect on gene expression or progeny production (Bürck et al., 2015). Nevertheless, the investigators showed that KAP1 was efficiently deSUMOylated, and thus inactivated as a repressor, during adenovirus infection seemingly through interactions with E1B55K (Bürck et al., 2015).

Based on these observations, we asked whether interactions between Ad5 proteins and KAP1 might provide the molecular switch necessary to initially release AAV from KAP1-mediated repression by targeting KAP1 for degradation, sequestration, or deSUMOylation, for example. To investigate this further, we monitored KAP1 protein levels in cells infected with increasing concentrations of either Ad5 or HSV-1, which revealed a dose-dependent degradation of KAP1 in the presence of either helper virus. Based on this observation, we propose that KAP1 targeting by Ad5 and HSV-1 represents an unidentified helper function for AAV replication necessary to relieve the viral genome from its latent state. These results then provoked the question of whether the modulation of KAP1 protein levels might be used to enhance recombinant AAV (rAAV) transduction or replication, as there is currently a strong drive to establish methods for the scaling-up of vector production and the design of vectors with greater bioactivity. Transduction and replication efficiency of various rAAV serotypes and vectors was

analyzed in the context of KAP1 depletion, revealing a 2-fold increase in rAAV replication in the presence of Ad5 coinfection.

6.2 Results

6.2.1 Ad5 and HSV-1 infection leads to the degradation of KAP1

Since basal levels of Rep expression under latent conditions are not sufficient to counteract KAP1, and depletion of KAP1 alone is not sufficient to trigger AAV2 transcription and replication, it is conceivable that a biological switch is necessary to allow for the upregulation of *rep* expression prior to the onset of p-KAP1-S824 accumulation. As AAV2 is a dependovirus, we hypothesized that helper viruses could potentially fulfill this role. Previous studies have shown that Ad5 interacts with the very close KAP1 (TIF1-β) relatives TIF1-α and TIF1-γ, resulting in the proteasome-mediated degradation of TIF1-γ as well as the reorganization of both TIF1-α and TIF1-γ to promyelocytic leukemia body tracks (Forrester et al., 2012; Yondola and Hearing, 2007). To investigate whether Ad5 targets KAP1 for degradation, we infected 293T and HeLa cells with increasing MOIs of Ad5 and monitored KAP1 levels by western blot. We observed a clear, dose-dependent depletion of KAP1, with almost complete loss at the highest MOI (Figure 38A). The reduction in KAP1 levels was not apparent in previous replication experiments as the low Ad5 MOI used in these experiments (2 PFU/cell) only resulted in minimal depletion of KAP1. In HeLa cells, we observed almost complete loss of KAP1 at an MOI of 5 PFU/cell (Figure 38B). KAP1 levels were rescued in the presence of 5μM MG132 in HeLa cells, suggesting that Ad5 targets KAP1 for proteasome-mediated degradation (Figure 38C). We next infected 293T cells with 10 PFU/cell of Ad5 and harvested at various time points to determine the kinetics of KAP1 degradation. At this MOI, degradation only became apparent at 48h post infection (38D). To determine whether this observation extended to other AAV2 helper viruses, we repeated these infection experiments using HSV-1. While the effect was less pronounced, HSV-1 infection also resulted in a 60% depletion of KAP1 over the range of MOIs tested for both 293T and HeLa cells (Figure 38E and F).

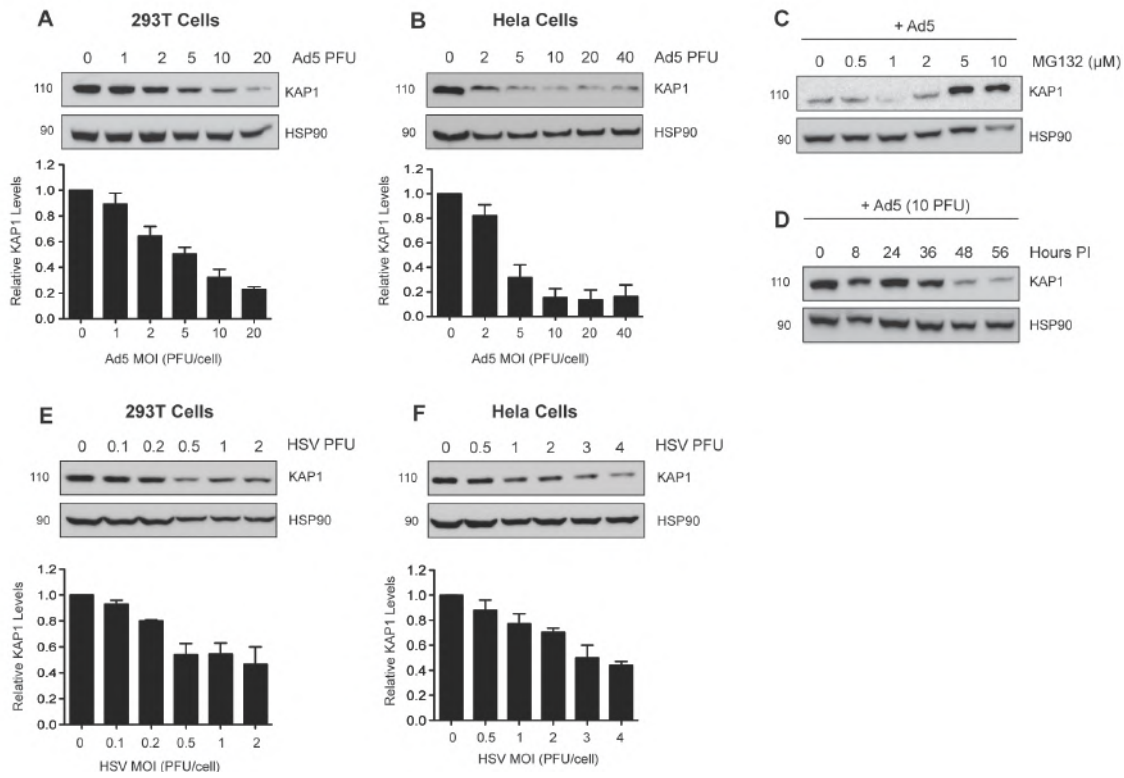


Figure 38. Ad5 and HSV-1 target KAP1 for degradation.
(A-B) Analysis of KAP1 levels by western blotting in 293T **(A)** and HeLa **(B)** cells 48h after infection with increasing concentrations (PFU/cell) of Ad5. **(C)** 293T cells pretreated with increasing concentration of the proteasome inhibitor MG132 were infected with 10 PFU/cell of Ad5. Cells were harvested 48h after infection for analysis of KAP1 levels by western blotting. **(D)** 293T cells infected with 10 PFU/cell Ad5 were harvested at the indicated time points for analysis of KAP1 levels by western blotting **(E-F)** Analysis of KAP1 levels by western blotting in 293T **(E)** and HeLa **(F)** cells 48h after infection with increasing concentrations (PFU/cell) of HSV-1.
Values are reported as mean±SEM, n=3.

6.2.2 KAP1 depletion does not enhance recombinant AAV transduction.

Recombinant AAV vectors (rAAV) have demonstrated great success as vectors for gene therapy in numerous clinical trials. However various challenges remain in bringing rAAV vectors more widely into the clinics. In particular, exceedingly high doses of rAAV are currently necessary to effectively target large human organs or the musculoskeletal system. This presents a real challenge at the level of vector production, and also increases the ordinarily low risk of complications related to vector immunogenicity. As a result, there is a strong drive to establish methods both

for the scaling-up of vector production, as well as for the design of vectors with greater bioactivity. We wondered if the manipulation of cellular KAP1 levels might be used to enhance either rAAV transduction or replication/production. This work suggests that KAP1-mediated repression of AAV2 is dependent on the presence of binding sites within the *rep* gene, which is not present in rAAV. However it is possible that further binding sites exist, within the viral ITRs for example, which could affect transduction efficiency. To address this, control and KAP1-depleted 293T cells were transduced with increasing MOIs of rAAV2, rAAV6, or rAAV9 expressing a CMV-GFP transgene cassette and were analyzed for GFP expression by FACS 48h post transduction. No difference in transduction efficiency was detectable between control and KAP1-depleted cells for any of the three serotypes tested (Figure 39A, B, and C). These results support the idea that the interaction between KAP1 and AAV2 is dependent on the presence of the *rep* gene.

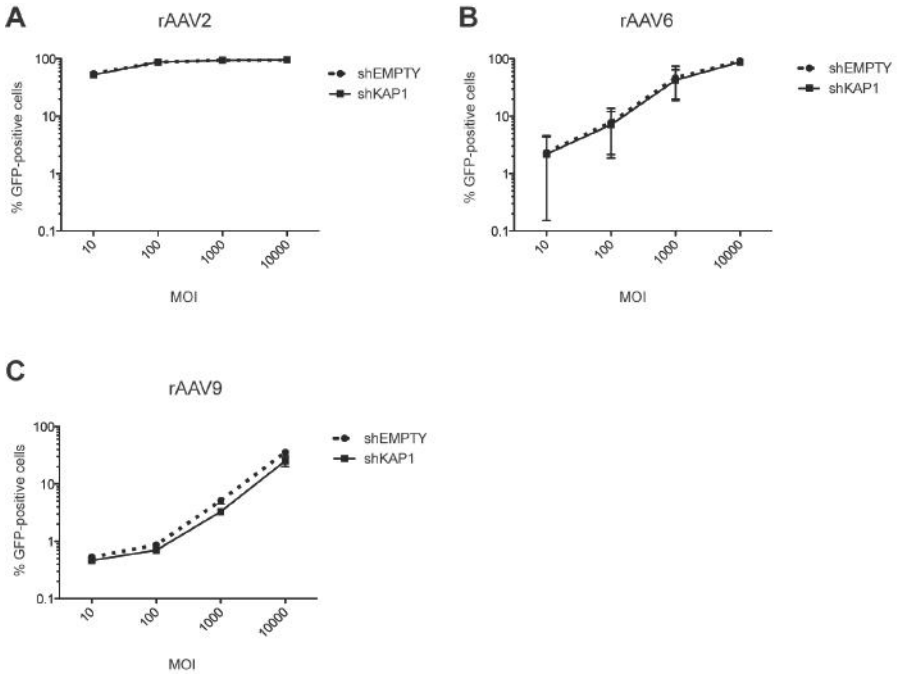


Figure 39. Depletion of KAP1 has no effect on recombinant AAV vector transduction. **(A-C)** 293T cells transduced with increasing concentrations (VG/cell) of **(A)** rAAV2-GFP, **(B)** rAAV6, or **(C)** rAAV9 were analyzed by FACS 48h after transduction for GFP expression. Values are reported as mean±SEM, n=3.

6.2.3 KAP1 depletion leads to modest enhancement in rAAV replication in the context of Ad5 infection.

We next considered whether rAAV production could be affected by KAP1 repression. Although provided in *trans,* helper plasmids containing the *rep* gene might be subject to KAP1-mediated transcriptional repression. To test this, control and KAP1-depleted cells were cotransfected with either a single-stranded (ss) or a self-complementary (sc) rAAV2-GFP vector, and the helper plasmid pDG, which contains *rep, cap,* and the minimal Ad5 helper functions necessary for AAV replication. Cells were harvested 72h post transfection for qPCR analysis of vector genomes. No difference was observed between control and KAP1-depleted cells transfected with either single-stranded or self-complementary rAAV2 (Figure 40A). We next performed replication experiments using Ad5 and mini-pDG, which contains only *rep* and *cap,* in place of the helper plasmid pDG. Control and KAP1-depleted 293T cells were transfected with a single-stranded rAAV2-GFP vector and mini-pDG and infected with Ad5 4h after transfection. Cells were harvested as described above. Replication in control cells was ~10-fold less efficient in the context of Ad5 infection than previously observed with pDG transfection (Figure 40B). Interestingly however, there was a 2-fold enhancement in rAAV2-GFP replication in KAP1-depleted cells as compared to controls in the presence of Ad5. These results suggest that KAP1 inhibits a helper function of Ad5 not present in pDG. While replication in the presence of pDG is clearly more efficient than with Ad5, it is possible that identifying and introducing this unknown helper function into pDG may further enhance rAAV replication.

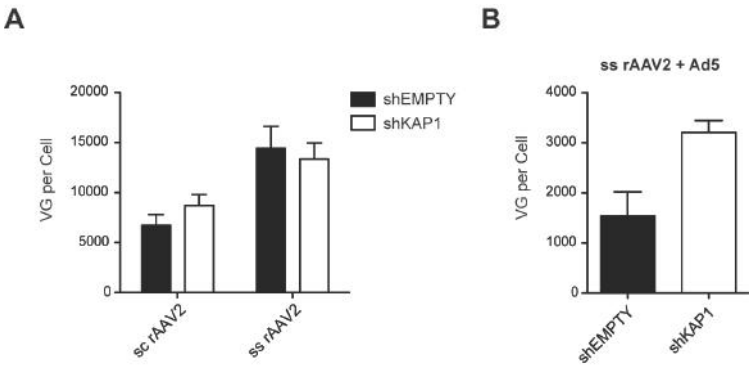


Figure 40. Recombinant AAV2 replication is modestly enhanced in KAP1-depleted cells in the presence of Ad5, but not when replication is supported by the helper plasmid pDG.

(A) Control (shEMPTY) or KAP1-depleted (shKAP1) 293T cells transfected with either a self-complementary (sc) or single-stranded (ss) rAAV2 vector and the helper plasmid pDG were harvested 72h after transfection for qPCR analysis of vector genomes. **(B)** Control (shEMPTY) or KAP1-depleted (shKAP1) 293T cells transfected with a single-stranded (ss) rAAV2 vector and the helper plasmid mini-pDG, containing only *rep* and *cap,* and infected with Ad5 (4 PFU/cell) 4h after transfection were harvested 72h after transfection for qPCR analysis of vector genomes.

Values are reported as mean±SEM, n=3.

6.3 Discussion

DNA viruses that replicate in the nucleus face the challenge of host cell chromatin modulation machinery, which controls access to the viral genome for the processes of transcription and DNA replication. The formation of nucleosomes on incoming viral genomes is necessary to create transcriptionally competent DNA, however it also provides a scaffold for the activity of cellular nucleosome- and histone-modifying proteins that can rapidly alter transcriptional activity from permissive to non-permissive, and vice versa.

The dynamic chromatin regulation of the HSV-1 genome is a critical regulatory determinant of both the latent and lytic phases. Upon nuclear release, the non-nucleosomal HSV-1 genome (Pignattii and Cassai, 1980) is rapidly assembled into chromatin (Conn and Schang, 2013; Lacasse and Schang, 2012) by one of two possible classes of histone chaperone complexes – the heterochromatin-associated Daxx/ATRX complex (Lukashchuk and Everett, 2010), or the euchromatin-associated HIRA/ASF1a complex (Oh et al., 2012) – and enrichment for either of these complexes in the particular subnuclear microenvironment into which the viral genome is released likely determines the progression of infection (Knipe and Cliffe, 2008). In the context of latency, the HSV-1 genome is quiescent, and nucleosomes associated with lytic genes bear repressive histone marks. For reactivation to take place, this repression is countered by the recruitment of various activating proteins – including TAF1-B/Set and the cellular coactivator complex, comprised of HCF-1 coupled to the histone demethylases LSD1 and JMJD2 – to viral IE promoters where they act together to shift the epigenetic balance to one of transcriptionally active euchromatin (Knipe and Cliffe, 2008).

Interestingly, the KAP1-interacting protein CHD3 was shown to be necessary for the initial repression of HSV-1 genomes (Arbuckle and Kristie, 2014). Repression of IE and E genes by CHD3 was observed as early as 30 minutes post infection, and the inhibition of HSV-1 gene expression that is associated with inhibition of the histone demethylase LSD1 was partially compensated for by CHD3 depletion. This suggests that CHD3 acts during

the initial repression of incoming viral genomes, a process that is normally circumvented by the HCF-1/LSD1 coactivator. The investigators in this study did not address how CHD3 is recruited to the HSV-1 genome however. It is well documented that KAP1 mediates the formation of repressive heterochromatin through the recruitment of CHD3 and the histone methyltransferase SETDB1, and so it is conceivable that HSV-1 targets KAP1 through an unknown mechanism in order to counteract the repressive effects of CHD3.

Alternatively, HSV-1 infection has been observed to result in the broad suppression of cellular protein synthesis through viral host shutoff mechanisms (Smiley, 2004). Host shutoff essentially stems from two distinct inhibitory pathways – (1) destabilization of existing mRNA during early infection by the HSV-1 virion host shutoff protein (vhs), a ribonuclease which enters the cell as a component of the viral tegument (Kwong and Frenkel, 1987), and (2) suppression of host mRNA synthesis by the multifunction IE protein ICP27 (Hardwick and Sandri-Goldin, 1994). This global suppression of cellular gene expression combined with HSV-1 manipulation of RNA pol II functions to optimize viral protein synthesis. Further analysis of the temporal changes in KAP1 transcription will be necessary to determine what contribution these processes make to the observed depletion of KAP1.

The adenovirus genome on the other hand enters the host nucleus already as a highly condensed nucleoprotein complex, which must be remodeled in order for transcription and replication to take place (Giberson et al., 2012). Similarly to HSV-1, the cellular proteins TAFI-B/Set, nucleophosmin, and the pp32 proteins work together to remodel the tightly compacted viral genome (Haruki et al., 2006; Spector, 2007; Xue et al., 2005), increasing accessibility of the viral DNA to transcriptional activators. KAP1 has been suggested to repress adenovirus replication at this initial stage by preventing the decondensation of the incoming viral genome (Bürck et al., 2015). Supporting this, we also observed a 2-fold enhancement in Ad5 replication in the context of KAP1 depletion (Chapter 4, Figure 23). It is possible then that Ad5 targets KAP1 for degradation to counteract this

repression, however the fact that we only observe degradation at late time points during infection makes this unlikely. One study has shown that KAP1 repressor activity is inhibited during adenovirus infection through interactions between Ad E1B55K and KAP1, which stimulate the deSUMOylation of KAP1 (Bürck et al., 2015). KAP1 has six lysine residues that are putative sites for SUMOylation (Lee et al., 2007). K554, K779, and K804 are the major sites shown to be necessary to support interactions with CHD3, and thus to mediate KAP1 repressive activity (Goodarzi et al., 2011; Lee et al., 2007). However, K767 SUMOylation in combination with S824 phosphorylation provides a signal for the targeted degradation of KAP1 by the SUMO-targeted E3 ubiquitin ligase, ring finger protein 4 (RNF4) (Kuo et al., 2014). It is possible that E1B55K leads to the selective deSUMOylation of K554, K779, and K804 to counteract KAP1 repression, but that SUMO-K767 combined with the phosphorylation of KAP1-S824 that we observed at ~42h post infection with Ad5 results in the observed degradation of KAP1 via RNF4. In other words, E1B55K-mediated deSUMOylation is sufficient to inhibit KAP1 with respect to adenovirus decondensation, and degradation is simply a secondary consequence of the altered state of KAP1 post-translational modifications. However, this provokes the question of why Rep-mediated accumulation of p-KAP1-S824 does not lead to a similar degradation. It will be interesting to dissect the particular SUMOylation states of KAP1 during Ad5, AAV2, and Ad5 + AAV2 infection. Equally, identifying the particular Ad5 proteins necessary to mediate this effect on KAP1 protein levels will be crucial to characterizing a mechanism. Supporting the mechanism outlined above, preliminary work in the lab suggests that transfection of E1B55K alone is sufficient to mediate KAP1 depletion.

In addition, we predict that during the early stages of coinfection – when helper virus replication is already ongoing but AAV2 remains latent – targeting of KAP1 by helper viruses is exploited by AAV2 for the release of its own genome from the latent state. In HeLa cells for example, considerable degradation of KAP1 was apparent even at the low MOIs equivalent to those used in previous replication experiments. This may help to explain why we

did not observe a strong effect of KAP1 depletion on AAV2 replication in HeLa cells, as Ad5-mediated degradation of KAP1 would have effectively rendered control and KAP1-depleted cells equivalent. Equally, E1B55K-mediated deSUMOylation of KAP1 in combination with E1A-mediated relief from YY1 repression of p5 may be sufficient for efficient *rep* expression without the need for KAP1 degradation. Why it is then necessary for AAV to also target KAP1 through phosphorylation might be explained by the complex interactions that take place between adenovirus and AAV during coinfection, which result in the inhibition of adenovirus gene expression by AAV (reviewed in section 1.3.3). As AAV replication progressively overpowers that of its helper virus, lytic replication is sustained through Rep-mediated phosphorylation of KAP1-S824.

These observations highlight the possibility that, although the minimal set of helper proteins for AAV replication have been identified for both Ad5 and HSV-1, the multiple roles of these factors in supporting AAV replication may be only partially understood. This provokes the question of whether there remain unknown cellular targets of helper virus factors, which might be modulated for the purpose of improving rAAV vector transduction and production. The observed 2-fold enhancement in rAAV replication in the presence of Ad5 suggests that KAP1 targets an adenovirus function that is irrelevant in the context of helper plasmid transfection, and furthermore that this function is beneficial to rAAV replication. For example, the process of efficient Ad5 decondensation might recruit factors involved in supporting AAV/rAAV replication, such as TAF1-B/Set and the ANP32 proteins, which have also been shown to be necessary for efficient AAV replication (Pegoraro et al., 2006). In addition, these factors may play a more important role under conditions in which AAV/rAAV replication is not optimally efficient, as observed in rAAV + Ad5 cells.

That no effect on rAAV replication was apparent in the presence of pDG alone despite the presence of a *rep* gene suggests that KAP1 is not being recruited to helper plasmids. Helper plasmids lack certain viral elements, such as the ITRs and the p5 promoter, which may be necessary for the recruitment of KAP1 through the formation of complex DNA

structures. For example, the chromatin-modifying protein HMG1 has been shown to promote the formation of Rep-DNA complexes (Costello et al., 1997); it is thought that HMG1-induced DNA bending produces allosteric transition structures that promote the recognition and binding of other proteins to form protein:DNA complexes. Additionally, HMG1 binds with high affinity to already distorted DNA structures (Pil and Lippard, 1992), such as ITRs. It is conceivable then that higher order DNA structures dependent on interactions between HMG1 and the viral ITRs are necessary for KAP1 recruitment. Alternatively, it is possible that the formation of DNA loops through interactions between KAP1-7SK snRNP complexes bound to p19 with YY1 bound to p5, which is not present in helper plasmids, are necessary to support the KAP1-*rep* interaction.

In conclusion, we show in this chapter that the AAV2 helper viruses Ad5 and HSV-1 both target KAP1, leading to its depletion apparently through proteasome-dependent degradation, although this will have to be further confirmed for both Ad5- and HSV-1-supported replication. This suggests the possibility that interference with KAP1 represents an unidentified helper function for AAV2 replication, and furthermore that KAP1 targeting is a conserved requirement for the efficient replication of DNA viruses. This led us to ask whether KAP1 might have a role in rAAV transduction or replication. While we were unable to demonstrate a role for KAP1 in rAAV transduction or replication in the presence of the helper plasmid pDG, a 2-fold enhancement in rAAV replication in the presence of Ad5 coinfection suggests that the interplay between Ad5, KAP1, and AAV/rAAV is more complex than anticipated. Importantly, it will be necessary to determine whether this observed enhancement in replication necessarily translates to increased rAAV particle production. Nevertheless, a more detailed understanding of the interactions involved may yet serve to further enhance rAAV replication and production in the context of pDG.

CHAPTER 7: GENERAL DISCUSSION AND FUTURE PERSPECTIVES

7.1 Introduction

AAV is remarkable in that it has evolved a unique, biphasic life cycle in which productive replication is dependent on both cellular host factors and coinfection by a helper virus, such as Ad5 or HSV-1. Unable to replicate autonomously however, infection by AAV alone leads to the establishment of latency either through long-term episomal persistence, or through preferential integration of the viral genome into specific sites in the human genome. These unique characteristics – the dependence on an unrelated virus for reactivation coupled with low immunogenicity and a capacity to establish latency through integration – are in notable contrast to what is known for other biphasic DNA viruses and have long fostered a view of AAV as somewhat of an anomaly.

The duality of the AAV lifestyle also necessitates an exceptional degree of complexity and flexibility from what is fundamentally a simple virus. AAV must be able to camouflage itself effectively against cellular defenses in order to persist during latency, while simultaneously remaining poised for rapid reactivation upon helper virus coinfection. Upon coinfection, AAV must regulate both its own replication and that of its helper virus such that optimal conditions for propagation are maintained. An interesting consequence of AAV's helper virus dependency is that both virus and host appear to benefit. It has been proposed that AAV has a protective effect on its host through the selective killing of cells infected with potentially pathogenic viruses, while remaining apparently innocuous in healthy cells. AAV may also benefit as the extra level of regulation imposed by helper dependency prevents its efficient replication from causing any deleterious effect to the host. In this respect, AAV might arguably represent a highly evolved virus – one that has successfully minimized any harmful effects on its host while simultaneously maximizing its pervasiveness – and thus presents a unique platform from which to study a novel class of host/pathogen/pathogen interactions.

The importance and breadth of the host/pathogen interactions necessary for completion of the complex AAV lifecycle are highlighted by the

striking simplicity of its genome. Having a limited repertoire of viral proteins, AAV must interact with an array of cellular factors to complete its life cycle. Indeed, a multitude of factors involved in transcription, RNA splicing, and the cellular DDR pathway have been shown to colocalize with AAV replication centers, as well as to copurify with Rep (Vogel et al., 2013). Although our understanding of the significance of such protein interactions in the AAV life cycle is constantly growing, there are still many roles of the astonishingly multifunctional Rep proteins that have yet to be fully explored.

The work presented in this thesis was intended to identify novel potential interaction partners for the AAV Rep proteins with the purpose of illuminating the processes involved in the regulation of AAV gene expression, and, furthermore, to define the mechanism of interaction. We initially set out to identify cellular factors involved in a mechanism of transcriptional repression by the Rep proteins that is dependent only on an intact NTP-binding site within Rep, and which has been observed for the AAV p5 promoter as well as various heterologous cellular and viral promoters (Dutheil et al., 2014; Kyöstiö et al., 1994). We performed BioID screens using each of the four Rep isoforms, which led to the identification of several interesting candidates involved in transcriptional regulation, including RDBP, RUVBL1, TCERG1, and KAP1. Physical interactions between Rep52, RDBP, RUVBL1, and KAP1 were confirmed by cross-linking IP, however no role for these candidates, or for TCERG1, could be demonstrated in the regulation of p5.

A role for KAP1 in the AAV2 lifecycle was discovered however when depletion of KAP1 led to a significant enhancement in AAV2 genome replication and transcription. Upon further analysis, we were able to demonstrate that KAP1 recruitment to the latent AAV2 genome triggers H3K9 trimethylation across the viral genome. Conversely, we show that the repressor activity of KAP1 is inactivated during lytic replication through phosphorylation of KAP1-S824 in a manner potentially dependent upon interactions between the Rep proteins, the phosphatase PP1, and its negative regulator NIPP1, and that these interactions may play a role in both AAV2 transcription and genome replication. Finally, based on the observation that Ad5 and HSV-1 both lead to the depletion of KAP1, we propose a model

in which latent AAV2 genomes are maintained in a repressed state by KAP1 until its partial degradation or manipulation by helper virus triggers the initial upregulation of *rep*. We envision that, during latency, KAP1 bound to the *rep* ORF serves to silence the downstream viral genome through histone methylation, while Rep proteins bound to AAV2 p5 maintain this essential early promoter free from repressive marks and thus competent for rapid reactivation (Figure 41A). Upon helper virus coinfection, previously described interactions between helper virus factors and cellular proteins bound to p5 upregulate *rep*, while interference with KAP1 by helper virus renders the genome permissive to transcription (Figure 41B and C). Newly expressed Rep proteins then form a complex with PP1 and NIPP1 and through these interactions inhibit the dephosphorylation of KAP1, leading to increased levels of p-KAP1-S824 and full reactivation of AAV2 (Figure 41D). It is also possible that modulation of KAP1 by Rep is necessary for the initiation of transcription of the “late” *cap* ORF.

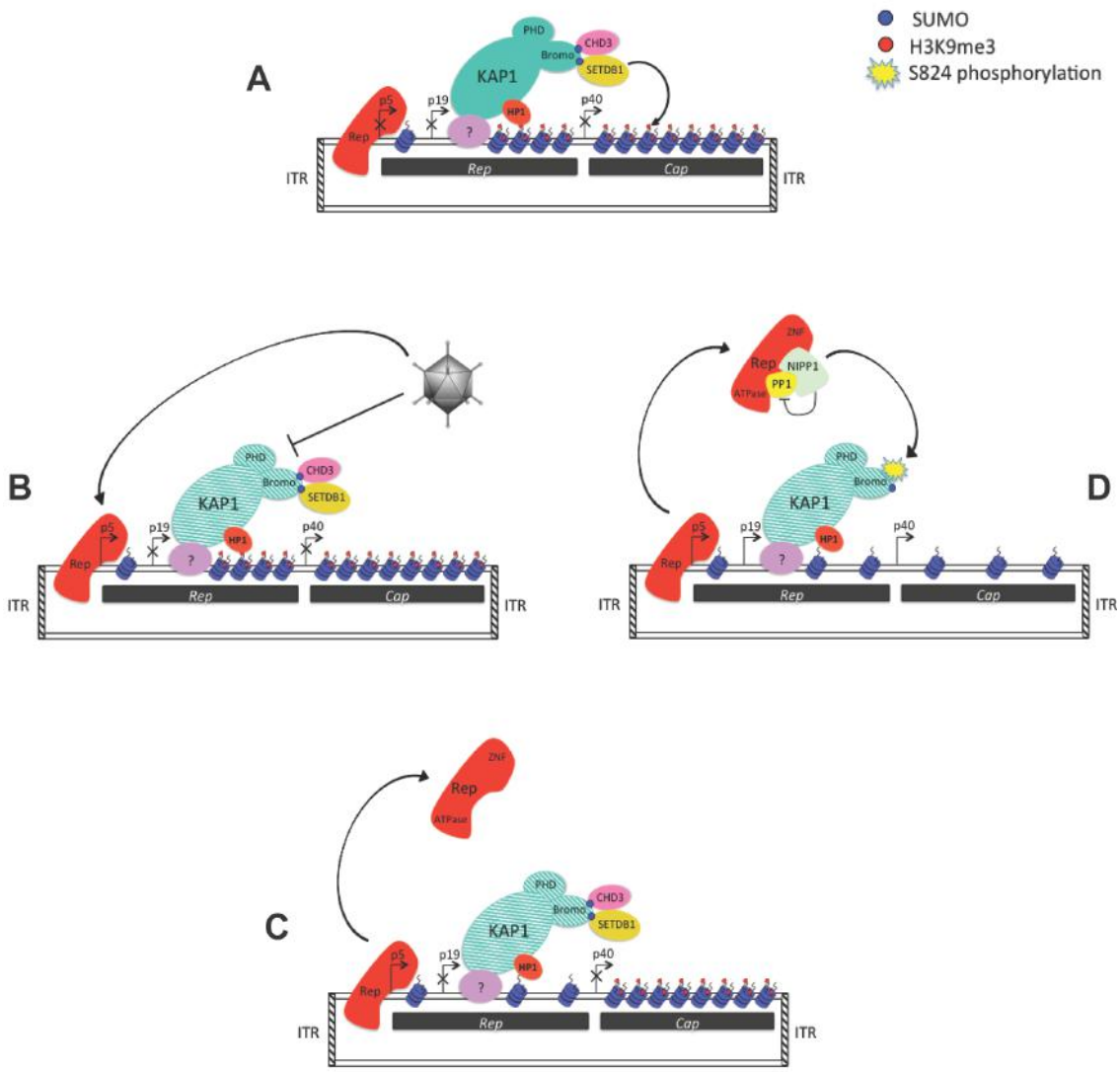


Figure 41. Model for release of AAV2 from KAP1-mediated latency.

(A) Incoming AAV2 genomes undergo second-strand synthesis, concatamerization, and chromatinization upon nuclear entry. KAP1 is recruited to the *rep* ORF via an unknown binding partner where it forms a scaffold for the recruitment of SETDB1 and CHD3, leading to the methylation of AAV2-associated histones. Rep bound to the p5 promoter both represses p5 activity while maintaining the promoter free from repressive marks. **(B)** Upon helper virus superinfection, repression of p5 is relieved through previously described interactions between helper virus proteins and cellular factors while interference with KAP1 by helper virus allows for (**C**) the initial upregulation of Rep. **(D)** Binding of Rep to PP1 and NIPP1 leads to enhanced levels of phosphorylated KAP1-S824 and release of the repressive complex.

7.2 Paving the way for a potential new field of AAV epigenetics

While all biphasic DNA viruses establish and maintain latency in a unique way, there still exist common themes that govern all persistent viral infections. Importantly, factors that restrict productive infection often positively and essentially affect the establishment of latency (Lieberman, 2016, 2008). For example, cellular sensing of incoming naked viral DNA leads to the formation of chromatin, which can repress the viral genome, but which is also necessary to protect the latent viral genome from DNA damage signaling and nucleolytic attack. Chromatinization of viral DNA also provides the scaffold required for the establishment of a dynamic transcriptional state, in which activating and repressive histone modifications are absolutely necessary for the correct temporal regulation of gene expression during the different stages of the biphasic viral life cycle. The rapidly expanding field of viral epigenetics has revealed an intricate web of interactions between viral and host cellular machinery where intrinsic cellular defense mechanisms are effectively coopted by the virus to promote viral latency. Various virally-encoded proteins can also modulate these repressive mechanisms – through the targeting of repressive chromatin assembly factors such as Daxx or ATRX, or through the recruitment of the histone demethylases JMJD2s and LSD1, for example – to regulate lytic and latent gene expression programs through the conversion of repressive heterochromatin to active euchromatin (Lieberman, 2016).

In line with its long-standing view as an outsider amongst DNA viruses, little is yet known about the nature and contribution of epigenetic marks to the genome organization and temporal gene regulation of AAV. Being a nuclear DNA virus however, it seems unlikely that AAV does not also encounter and interact with similar cellular defenses as do other DNA viruses, thus becoming subject to the epigenetic control of its genome. Given the importance of these interactions in the life cycles of DNA viruses such as HSV, Ad, EBV, HPV, CMV, and KSHV, it feels imperative that we develop a deeper understanding of these phenomena for AAV, not only to understand how AAV is affected by epigenetic modifications to its genome, but also how it may counteract these effects itself, or hijack other viral proteins to do so.

The findings presented in this work may thus open the door to a new field of AAV study focused on epigenetics. We show here that the latent AAV2 genome, with the exception of the p5 promoter, is enriched for the repressive mark H3K9me3. There is an ever-expanding list of possible histone modifications however, each involved in a considerable degree of crosstalk with one another (Bannister and Kouzarides, 2011), which could have a profound effect on the overall transcriptional state of the viral genome. It will therefore be necessary to determine the global epigenetic landscape of the AAV genome throughout latency and lytic replication and to relate changes in gene expression to changes in chromatin states. Identifying the particular modifications involved may also lead to the identification of new cellular proteins involved in the modification of AAV chromatin, and furthermore may reveal novel potential targets of the Rep proteins for the counteraction, or domestication, of these activities.

These findings also highlight the possibility that AAV is not so fundamentally dissimilar from other DNA viruses and that certain phenomena observed for other viruses might be extrapolated to AAV. Equally, the absence of such phenomena in AAV might represent a key factor in the difference between autonomy and helper dependency. One question that remains to be answered is whether AAV establishes latency through integration or episomal persistence *in vivo.* Viruses that establish latency in proliferating cells must have a mechanism in place to maintain their genomes. For viruses that do not integrate, this involves tethering of the viral DNA to metaphase chromosomes by virally encoded DNA-binding proteins, an activity that has never been demonstrated for the Rep proteins. Interestingly, one of the proteins identified in our BioID screen, BRD4, is a chromatin-associated factor involved in the E2-mediated chromosome tethering of HPV (Abbate et al., 2006) and the targeting of MLV integration through tethering of viral genomes to regions of polyacetylated histones (DeRijck et al., 2013), suggesting that the possibility of AAV tethering may require further investigation. Alternatively, a lack of tethering activity for AAV would certainly explain the need to evolve a mechanism for integration,

particularly in proliferating tissues. The ability to establish latency by two different mechanisms may be yet another example of how AAV is able to be so promiscuous in its tissue tropism. As mentioned previously, both HSV-1 and Ad5 play highly active roles in the modulation of their chromatin states upon entry and reactivation. It will be interesting therefore to determine which chromatin-modifying functions, if any, of Ad5 and HSV-1 proteins might be necessary for AAV replication. It is also tempting to speculate that the loss of some autonomous chromatin-modifying activity by AAV might have rendered it helper-dependent.

To date, the only tissue found to have high frequencies of AAV DNA is muscle, where 17% of random biopsies tested positive (Tezak et al., 2000). Most viruses establish latency only in specialized cell types where conditions are ideal for enabling viral latency. For example, predominant latent forms of EBV are found in CD19+ memory B cells, and CMV in CD34+ HSCs. In the case of CMV, CD34+ HSCs in effect force viral latency through the concerted effect of a lack of viral activators and the presence of latency-associated repressors, such as KAP1, which act together to establish a repressive heterochromatin state on the viral genome (reviewed in Poole and Sinclair 2015). Following differentiation into macrophages or dendritic cells, changes in the nuclear environment, including the loss of KAP1 expression, lead to the chromatin-mediated reactivation of lytic CMV replication. In this scenario, viral latency appears to be almost completely at the mercy of the nuclear microenvironment and its influence on chromatin states. It is conceivable then that defining the particular set of chromatin-modifying proteins involved in establishing AAV latency will help to guide our search for the latent reservoir and possibly illuminate the deciding factors in the choice between integration and episomal persistence. It would be interesting, for example, to make the distinction of whether AAV sequences detected in muscle biopsies are derived only from terminally differentiated muscle cells, from tissue-resident muscle stem cells (satellite cells), or from both, and to relate these differences to changes in nuclear microenvironment, epigenetic programming, and proliferative potential.

7.3 PP1 targeting as a novel mechanism for the inactivation of KAP1 corepressor activity

One characteristic of KSHV reactivation is the vPK-mediated inactivation of KAP1 through phosphorylation of serine 824 (Chang et al., 2009), presumably necessary to fully relieve the viral genome from KAP1-mediated repression. Here, we show that lytic replication of AAV in the presence of Ad5 is also strongly associated with a phosphorylation of KAP1. Furthermore, we show that this phosphorylation is the direct result of an AAV countermeasure to KAP1 through the Rep-mediated inhibition of PP1. Interestingly, several viruses have already been shown to target PP1 for a variety of reasons. Upon viral infection, the protein kinase PKR is activated and leads to the phosphorylation of the initiation factor eIF2α, resulting in the global inhibition of protein synthesis and halting viral replication. In order to counteract this, viruses have found ways to prevent the accumulation of phosphorylated eIF2α. The HSV-1 protein ICP34.5 leads to the specific dephosphorylation of eIF2α by forming a bridge between it and PP1, effectively bypassing infection-mediated ribosomal shutdown (Li et al., 2011). The HPV E6 protein achieves the same goal by interacting with the GADD34/PP1 complex (Kazemi et al., 2004). PP1 is also involved in the regulation of CREB signaling through the dephosphorylation and inactivation of CREB proteins, which consequently results in a down-regulation of CRE-regulated elements (Bito et al., 1996; Hagiwara et al., 1992). Both EBV and HBV have been shown to target PP1 for this reason. One of the key proteins in EBV-mediated immortalization is the CRE-regulated latent membrane protein 1 (LMP1), and it is thought that the viral protein EBNA2 inhibits PP1 in order to increase expression of this essential factor (Fahraeus et al., 1994). Similarly, inhibition of PP1 by the HBV regulatory protein HBx leads to enhanced CRE-regulated viral transcription (Cougot et al., 2012). Finally, the V proteins of the measles virus and the related paramyxovirus Nipah virus interact with PP1, preventing PP1-mediated dephosphorylation of the RNA

sensor MDA5 and allowing for evasion of innate immune recognition (Davis et al., 2014).

Here, we show that the AAV Rep52 and Rep78 proteins also inhibit PP1 activity, possibly by bridging the interaction between PP1 and the negative regulator NIPP1, and that this inhibition leads to increased levels of phosphorylated, inactive KAP1. Importantly, we demonstrate that Rep proteins unable to interact with PP1 are also unable to support both viral transcription and DNA replication, providing an essential link between Rep-mediated inhibition of PP1, phosphorylation of KAP1, and AAV replication. This is not only the first example of PP1 targeting by a parvovirus, but also of PP1 targeting for the purpose of regulating KAP1 activity. Additionally, the hypothesis that Rep manipulates PP1 activity by bridging its interaction with a negative regulator represents a novel mechanism of viral PP1 targeting. Alternatively, Rep might behave in a similar manner to HPV E6, selectively interacting with and recruiting NIPP1/PP1 complexes to KAP1 and effectively competing out the binding of active PP1.

While KSHV has been shown to target KAP1 directly via the viral kinase vPK, there is evidence to suggest that an as yet unidentified CMV-encoded protein also targets KAP1 for phosphorylation (Rauwel et al., 2015). It will be interesting to determine if either KSHV or CMV uses a similar mechanism to AAV in the relief of their genomes from KAP1-mediated repression. Furthermore, it would be interesting to determine whether this mechanism extends to other members of the parvovirus family. Initial analysis of the sequences of both B19 and MVM NS1 proteins revealed potential PP1 consensus binding sites. Although it is not fully understood whether the parvovirus B19 establishes latency as a rule, reactivation of latent virus in immunocompromised patients can lead to persistent and relapsing viremia, suggesting the presence of a latent reservoir. It is tempting to speculate that KAP1 might play a similar role for B19.

7.4 Significance for AAV gene therapy

AAV gene therapy is demonstrating great success in multiple ongoing clinical trials, highlighting the enormous potential for this new experimental therapy in the treatment of debilitating monogenic diseases. Various challenges remain in bringing rAAV vectors into the clinics however. In particular, high doses of rAAV are currently necessary to effectively target large human organ systems. This presents a real challenge at the level of vector production, and also increases the ordinarily low risk of complications related to vector immunogenicity. As a result, there is a strong drive to establish methods both for the scaling-up of vector production, as well as for the design of vectors with greater bioactivity. Significant enhancements in AAV vector design have previously been derived from an enhanced understanding of wild type AAV2 biology. Given that the biology of rAAV vectors likely mimics that of latent viral genomes, the findings presented here have particular relevance for gene therapy with strong implications for the improvement of AAV vector design.

Although we were unable to demonstrate a role for KAP1 in rAAV transduction or replication in the presence of the helper plasmid pDG, the moderate effect observed in the presence of minipDG and Ad5 coinfection highlights the need to further investigate the possibility that KAP1 depletion may improve vector production. It is possible that proteins recruited during the process of Ad decondensation may be beneficial for the initial reactivation of AAV, and that the activating properties of such proteins are effectively counteracted by KAP1. In defining the potential role for Ad5-recruited chromatin-modifying proteins in AAV replication, we may uncover new targets, as well as new host factors, that might be manipulated for the purposes of vector production. Preliminary data in the lab have revealed that Ad5 E1B55K may be necessary for the degradation of KAP1. Given that 293T cells have been immortalized with Ad E1/E2 genes, these observations also provoke the more practical question of whether 293T cells truly represent the ideal cell line in which to produce rAAV. We show that KAP1 degradation by Ad5 is significantly more pronounced in HeLa cells than in 293T cells. This might be representative of the development in 293T cells of (1) compensatory mechanisms to combat E1B55K-mediated degradation of

KAP1, and (2) redundant mechanisms to compensate for KAP1 depletion. Both possibilities would have implications for rAAV replication in 293T cells.

Interestingly, rAAV-expressed transgene expression has not yet been shown to be affected by either DNA or histone methylation (Dupont et al., 2015; Léger et al., 2011; Penaud-Budloo et al., 2008). Instead, diminished transgene persistence has been related to such factors as a progressive loss of therapeutic vector genomes over time, and increased levels of transgene mRNA oxidation (Dupont et al., 2015; Le Hir et al., 2013; Manno et al., 2006). These studies were carried out on a limited number of transgenes and target tissues however. Given the rapid expansion of AAV gene therapy and its application for an ever-growing number of diseases, it may still prove beneficial to screen new desired transgenes and promoters for sites of KAP1 or KRAB-ZFP binding, the modification of which would avoid unwanted transcriptional silencing.

Another challenge in gene therapy is balancing efficient *in vivo* transduction in susceptible tissues while exclusively targeting a desired organ. Current strategies for achieving this include the use of tissue-specific promoters and enhancers to drive expression of the transgene in question, and packaging of the recombinant vector genome within an AAV serotype having the desired tissue tropism. The various AAV serotypes are not specific for one target tissue however, and instead target different subsets of tissues; similarly, strong promoter/enhancer elements have the potential to generate leaky expression in undesired tissues. One elegant strategy to achieve specific transgene expression involves the exploitation of post-transcriptional regulation mechanisms to degrade transgene mRNA through the use of endogenous tissue-specific microRNA (miRNA) systems. Specifically, the incorporation of target sequences for the liver-specific miR-122T into the 3'-UTR of an AAV vector cassette was successful in inhibiting AAV9-mediated gene expression in the liver (Qiao et al., 2011). In situations where a suitable tissue-specific miRNA does not exist, the incorporation of target sites for tissue-specific KRAB-ZFPs might represent another alternative, leading to selective KAP1-mediated repression of vector genomes.

7.5 Future Perspectives

Bearing in mind the layers of complexity that are created through the amalgamation of AAV's helper dependency and its susceptibility to epigenetic mechanisms, it is clear that the work presented here represents merely the tip of the iceberg. There are still many questions left to answer in order to fully define the mechanism and consequence of KAP1 activity on AAV. One of the first will be to identify the recruitment factor for KAP1. One possibility is that KAP1 is recruited by YY1, as the two proteins have been shown to interact with each other in mouse ES cells. This can be verified using KAP1-specific ChIP experiments in latently infected control and YY1-depleted cells to determine whether KAP1 is still recruited in the absence of YY1. In the event that YY1 is not the factor responsible, potential proteins involved might be purified using a combination of EMSA followed by mass spectrometry analysis. By comparing shift patterns of latent AAV2 fractionated from control or KAP1-depleted cells, it may be possible to specifically purify a band representing AAV2 genomes bound by the KAP1 repressive complex. Purification of this band followed by mass spectrometry analysis of bound proteins would potentially identify new candidate proteins necessary for KAP1 recruitment. Alternatively, lysates from a KAP1-specific cross-linked IP might be separated by EMSA and probed for AAV sequences to the same end. Answering this question will not only be instrumental in further defining the mechanism of KAP1 action, but will also have implications for tissue-specific changes in AAV behavior, as KAP1 recruitment factors often have tissue-specific expression profiles.

KAP1 recruitment factors are generally considered to be KRAB-ZFPs, which interact with KAP1 via its RBCC domain. However, KAP1 deleted for this domain is still recruited to numerous sites in the human genome. Our cross-linked IP experiments demonstrate that the minimal Rep ATPase domain is sufficient to mediate interactions between Rep and KAP1, however we have yet to identify which domains of KAP1 are involved. This can be determined by CL-IP using mutants of KAP1 deleted for the various domains. Additionally, it may be possible to determine which domains are necessary for recruitment of KAP1 to the AAV2 genome by performing KAP1-specific ChIP in cells depleted for endogenous KAP1 and reconstituted with the

various deletion mutants. These results might be very helpful not only in identifying a recruitment factor for KAP1, but also in understanding the role for KAP1 in binding AAV2. For example, while binding of KAP1 to 3' coding exons of ZNF genes and H3K9me3 deposition are associated with RBCC-dependent recruitment, binding of KAP1 to TSS and associated proximal promoter pausing of RNA pol II is not. In parallel to these experiments, it will be interesting to analyze the effect of Rep expression on KAP1 protein interactions, or during latent versus lytic replication, as this could help to detect any potential changes in function of KAP1 in the context of AAV.

Furthermore, it will also be necessary to fully characterize the dynamics of epigenetic changes that take place on the AAV genome throughout the lifecycle. ChIP experiments for the various activating and repressive histone modifications will be assessed under latent conditions. It may be possible to circumvent the problems that we encountered previously when attempting to perform H3K9me3-specific ChIP on actively replicating AAV2 by harvesting cells at much earlier time points, which would avoid the presence of high levels of AAV2 DNA. The challenge will be in determining the ideal time-point at which changes in AAV2-associated H3K9 methylation have taken place, but DNA replication has not yet begun in earnest. Alternatively, it might be possible to generate replication defective AAV2 particles bearing the Rep-K340H mutation by providing Rep *in trans*, which could then be used to compare +/- Ad5 conditions without any active replication. The risk in this approach would lie in the fact that WT Rep may be necessary for the recruitment of KAP1 to the AAV genome. In addition to histone modifications, it will be interesting to further analyze the involvement of other members of the KAP1 repressive complex, such as HP1. Taking a broader view, interactions between Rep and components of several other chromatin-modifying complexes were detected in our screen – including the CAF1-A, PTW/PP1, and HCF-1 complexes – which will require further investigation.

The studies described above will not only provide us with an understanding of the dynamic AAV epigenetic landscape, but will also complement a deeper analysis of the effect of KAP1 on base-line AAV transcription. Up to this point, we have yet to directly demonstrate an effect of

KAP1 depletion on AAV transcription. The requirement for helper factors to initiate transcription is once more a complicating factor; without them, transcription is very inefficient, while in their presence, transcription is so efficient as to mask any early differences. It may again be beneficial to try much earlier time points – within 30 minutes to an hour after infection for example – when AAV2 in KAP1-depleted cells may still have an advantage. In parallel to these and the above outlined work, it will be interesting to analyze RNA pol II occupancy across the viral genome in the presence and absence of KAP1 and to correlate these changes with KAP1 occupancy, histone modifications, and changes in promoter transcriptional activity.

In conclusion, we demonstrate for the first time that AAV2 latency is regulated through the epigenetic modification of its genome by the corepressor KAP1. This challenges the long-standing model whereby latent AAV2 is silenced solely through the binding of its genome by the cellular factors YY1 and MLTF, and the viral Rep proteins. We provide evidence that the helper viruses Ad5 and HSV-1 target KAP1 for degradation and suggest that this not only represents an unknown helper function for AAV2 replication, but also a conserved requirement for the efficient replication of DNA viruses. Furthermore, we characterize a novel mechanism of PP1 antagonism through which the AAV2 Rep proteins further counteract KAP1-mediated repression, thus establishing an active role for AAV2 in the regulation of its epigenetic landscape. This work provides the first steps towards the development of a new field of AAV biology focused on the inevitable contribution of chromatin dynamics, with the potential for significant innovation in the field of AAV gene therapy.

BIBLIOGRAPHY

Abbate, E.A., Voitenleitner, C., and Botchan, M.R. (2006). Structure of the Papillomavirus DNA-Tethering Complex E2:Brd4 and a Peptide that Ablates HPV Chromosomal Association. Mol. Cell *24*, 877–889.

Abrink, M., Ortiz, J.A., Mark, C., Sanchez, C., Looman, C., Hellman, L., Chambon, P., and Losson, R. (2001). Conserved interaction between distinct Kruppel-associated box domains and the transcriptional intermediary factor 1 beta. Proc Natl Acad Sci U S A *98*, 1422–1426.

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat. Rev. Genet. *13*, 720–731.

Agbandje-McKenna, M., and Kleinschmidt, J. (2011). AAV Capsid Structure and Cell Interactions. In Adeno-Associated Virus, pp. 47–92.

Agbandje-McKenna, M., Llamas-Saiz, a L., Wang, F., Tattersall, P., and Rossmann, M.G. (1998). Functional implications of the structure of the murine parvovirus, minute virus of mice. Structure *6*, 1369–1381.

Akgün, E., Ziegler, M., and Grez, M. (1991). Determinants of retrovirus gene expression in embryonal carcinoma cells. J. Virol. *65*, 382–388.

Alazard-Dany, N., Nicolas, A., Ploquin, A., Strasser, R., Greco, A., Epstein, A.L., Fraefel, C., and Salvetti, A. (2009). Definition of herpes simplex virus type 1 helper activities for adeno-associated virus early replication events. PLoS Pathog. *5*, 1–12.

Allouch, A., Di Primio, C., Alpi, E., Lusic, M., Arosio, D., Giacca, M., and Cereseto, A. (2011). The TRIM family protein KAP1 inhibits HIV-1 integration. Cell Host Microbe *9*, 484–495.

Amelio, A., McAnany, P., and Bloom, D. (2006). A chromatin insulator-like element in the herpes simplex virus type 1 latency-associated transcript region binds CCCTC-binding factor and displays enhancer-blocking and silencing activities. J Virol *80*, 2358–2368.

Ansari, M.A., Singh, V.V., Dutta, S., Veettil, M.V., Dutta, D., Chikoti, L., Lu, J., Everly, D., and Chandran, B. (2013). Constitutive interferon-inducible protein 16-inflammasome activation during Epstein-Barr virus latency I, II, and III in B and epithelial cells. J. Virol. *87*, 8606–8623.

Arbuckle, J.H., and Kristie, T.M. (2014). Epigenetic repression of herpes simplex virus infection by the nucleosome remodeler CHD3. MBio *5*, e01027-13.

Atchison, R.W., Casto, B.C., Hammon, M., and Hammon, W.M. (1965). Adenovirus-associated defective virus particles. Science (80-.). *149*, 754–756.

Aydemir, F., Salganik, M., Resztak, J., Singh, J., Bennett, A., Agbandje-McKenna, M., and Muzyczka, N. (2016). Mutants at the 2-fold interface of AAV2 structural proteins suggest a role in viral transcription for AAV capsids. J Virol *90*.

Ayyanathan, K., Ayyanathan, K., Lechner, M.S., Lechner, M.S., Bell, P., Bell, P., Maul, G.G., Maul, G.G., Schultz, D.C., and Schultz, D.C. (2003). Heterochromatin protein 1 (HP1) is a key component of constitutive heterochromatin in. Genes Dev. 1855–1869.

Baker, A., Rohleder, K.J., Hanakahi, L. a, and Ketner, G. (2007). Adenovirus E4 34k and E1b 55k oncoproteins target host DNA ligase IV for proteasomal degradation. J. Virol. *81*, 7034–7040.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Bantel-Schaal, U., and Hausen, H. (1984). Characterization of the DNA of a Defective Human Parvovirus Isolated from a Genital Site. Virology *134*, 52–63.

Bantel-Schaal, U., and Zur Hausen, H. (1988). Adeno-associated viruses inhibit SV40 DNA amplification and replication of herpes simplex virus in SV40-transformed hamster cells. Virology *164*, 64–74.

Bantel-Schaal, U., Hub, B., and Kartenbeck, J. (2002). Endocytosis of adeno-associated virus type 5 leads to accumulation of virus particles in the Golgi compartment. J. Virol. *76*, 2340–2349.

Bardelli, M., Zárate-Pérez, F., Agúndez, L., Linden, R.M., Escalante, C.R., and Henckaerts, E. (2016). Identification of a functionally relevant AAV Rep68 oligomeric interface. J. Virol. JVI.00356-16.

Barklis, E., Mulligan, R.C., and Jaenisch, R. (1986). Chromosomal position or virus mutation permits retrovirus expression in embryonal carcinoma cells. Cell *47*, 391–399.

Bartlett, J.S., Wilcher, R., and Samulski, R.J. (2000). Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors. J. Virol. *74*, 2777–2785.

Batchu, R., and Hermonat, P. (1995). The trans-inhibitory Rep78 protein of adeno-associated virus binds to TAR region DNA of the human immunodeficiency virus type 1 long terminal repeat. FEBS Lett *367*, 267–271.

Batchu, R.B., Shammas, M.A., Wang, J.Y., and Munshi, N.C. (2001). Dual level inhibition of E2F-1 activity by adeno-associated virus Rep78. J Biol Chem *276*, 24315–24322.

Beaton, A., Palumbo, O., and Berns, K. (1989). Expression from the adeno- associated virus p5 and p19 promoters is negatively regulated in trans by the rep protein. J Virol *63*, 4450–4454.

Becerra, S.P., Koczot, F., Fabisch, P., and Rose, J. a (1988). Synthesis of adeno-associated virus structural proteins requires both alternative mRNA splicing and alternative initiations from a single transcript. J. Virol. *62*, 2745–2754.

Beer, D.G., Kardia, S.L.R., Huang, C.-C., Giordano, T.J., Levin, A.M., Misek, D.E., Lin, L., Chen, G., Gharib, T.G., Thomas, D.G., et al. (2002). Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat. Med. *8*, 816–824.

Berns, K., and Rose, J. (1970). Evidence for a single-stranded adenovirus-assoicated virus genome: Isolation and separation of compelementary single strands. J Virol *5*, 693–699.

Berns, L., Pinkerton, T., Thomas, G., and Hoggan, M. (1975). Detection of adeno-associated virus (AAV)-specific nucleotide sequences in DNA isolated from latently infected Detroit 6 cells. Virology *68*, 556–560.

Berthet, C., Raj, K., Saudan, P., and Beard, P. (2005). How adeno-associated virus Rep78 protein arrests cells completely in S phase. Proc Natl Acad Sci U S A *102*, 13634–13639.

Beullens, M., Van Eynde, A., Stalmans, W., and Bollen, M. (1992). The isolation of novel inhibitory polypeptides of protein phosphatase 1 from bovine thymus nuclei. J. Biol. Chem. *267*, 16538–16544.

Beullens, M., Vulsteke, V., Van Eynde, A., Jagiello, I., Stalmans, W., and Bollen, M. (2000). The C-terminus of NIPP1 (nuclear inhibitor of protein phosphatase-1) contains a novel binding site for protein phosphatase-1 that is controlled by tyrosine phosphorylation and RNA binding. Biochem J *352 Pt 3*, 651–658.

Bevington, J., Needham, P., Verrill, K., Collaco, R., Basrur, V., and Trempe, J. (2007).

Adeno-associated virus interactions with B23/Nucleophosmin: Identification of sub-nucleolar virion regions. Virology *357*, 102–113.

Bito, H., Deisseroth, K., and Tsien, R.W. (1996). CREB phosphorylation and dephosphorylation: A Ca2+- and stimulus duration-dependent switch for hippocampal gene expression. Cell *87*, 1203–1214.

Blacklow, N., Hoggan, M., and Rowe, W. (1967). Isolation of adenovirus-associated viruses from man. Proc Natl Acad Sci U S A *58*, 1410–1415.

Blacklow, N., Hoggan, M., Kapikian, A., Austin, J., and Rowe, W. (1968). Epidemiology of adenovirus-associated virus infection in a nursery population. Epidemiology *88*, 368–378.

Blacklow, N., Hoggan, M., Sereno, M., Brandt, C., Kim, H., Parrott, R., and Chanock, R. (1971). A seroepidemiologic study of adenovirus-associated virus infection in infants and children. Am. J. Epidemiol. *94*, 359–366.

Blanchette, P., Cheng, C.Y., Yan, Q., Ketner, G., Ornelles, D. a, Dobner, T., Conaway, R.C., Conaway, J.W., and Branton, P.E. (2004). Both BC-box motifs of adenovirus protein E4orf6 are required to efficiently assemble an E3 ligase complex that degrades p53. Mol. Cell. Biol. *24*, 9619–9629.

Bleker, S., Pawlita, M., and Kleinschmidt, J. a (2006). Impact of capsid conformation and Rep-capsid interactions on adeno-associated virus type 2 genome packaging. J. Virol. *80*, 810–820.

Bohenzky, R.A., and Berns, K.I. (1989). Interactions between the termini of adeno-associated virus DNA. J. Mol. Biol. *206*, 91–100.

Borden, K.L. (1998). RING fingers and B-boxes: zinc-binding protein-protein interaction domains. Biochem. Cell Biol. *76*, 351–358.

Brister, J.R., and Muzyczka, N. (1999). Rep-mediated nicking of the adeno-associated virus origin requires two biochemical activities, DNA helicase activity and transesterification. J Virol *73*, 9325–9336.

Brown, K., and Gilmartin, G. (2003). A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. Mol Cell *12*.

Bunch, H., Zheng, X., Burkholder, A., Dillon, S.T., Motola, S., Birrane, G., Ebmeier, C.C., Levine, S., Fargo, D., Hu, G., et al. (2014). TRIM28 regulates RNA polymerase II promoter-proximal pausing and pause release. Nat Struct Mol Biol *21*, 876–883.

Bunch, H., Zheng, X., Burkholder, A., Dillon, S.T., Birrane, G., Ebmeier, C.C., Levine, S., Fargo, D., Taatjes, D.J., and Calderwood, S.K. (2015). and pause release. *21*, 876–883.
Bürck, C., Mund, A., Berscheminski, J., Kieweg, L., Müncheberg, S., Dobner, T., and Schreiner, S. (2015). KAP1 is a host restriction factor that promotes HAdV E1B-55K SUMO modification. J. Virol. *90*, JVI.01836-15.

Cai, Y., Jin, J., Yao, T., Gottschalk, A.J., Swanson, S.K., Wu, S., Shi, Y., Washburn, M.P., Florens, L., Conaway, R.C., et al. (2007). YY1 functions with INO80 to activate transcription. Nat. Struct. Mol. Biol. *14*, 872–874.

Calcedo, R., Vandenberghe, L.H., Gao, G., Lin, J., and Wilson, J.M. (2009). Worldwide epidemiology of neutralizing antibodies to adeno-associated viruses. J Infect Dis *199*, 381–390.

Cammas, F., Mark, M., Dollé, P., Dierich, A., Chambon, P., and Losson, R. (2000). Mice lacking the transcriptional corepressor TIF1beta are defective in early postimplantation development. Development *127*, 2955–2963.

Cammas, F., Oulad-Abdelghani, M., Vonesch, J.-L., Huss-Garcia, Y., Chambon, P., and Losson, R. (2002). Cell differentiation induces TIF1beta association with centromeric heterochromatin via an HP1 interaction. J. Cell Sci. *115*, 3439–3448.

Cammas, F., Herzog, M., Lerouge, T., Chambon, P., and Losson, R. (2004). Association of the transcriptional corepressor TIF1β with heterochromatin protein 1 (HP1): An essential role for progression through differentiation. Genes Dev. *18*, 2147–2160.

Capili, A.D., Schultz, D.C., Rauscher, F.J., and Borden, K.L.B. (2001). Solution structure of the PHD domain from the KAP-1 corepressor: Structural determinants for PHD, RING and LIM zinc-binding domains. EMBO J. *20*, 165–177.

Carreira, A., Menendez, M., Reguera, J., Almendral, J.M., and Mateu, M.G. (2004). In Vitro Disassembly of a Parvovirus Capsid and Effect on Capsid Stability of Heterologous Peptide Insertions in Surface Loops. J. Biol. Chem. *279*, 6517–6525.

Carson, C.T., Orazio, N.I., Lee, D. V, Suh, J., Bekker-Jensen, S., Araujo, F.D., Lakdawala, S.S., Lilley, C.E., Bartek, J., Lukas, J., et al. (2009). Mislocalization of the MRN complex prevents ATR signaling during adenovirus infection. EMBO J. *28*, 652–662.

Carter, B., Khoury, G., and Rose, J. (1972). Adenovirus-Associated Virus Multiplication. J Virol *10*, 1118–1125.

Carter, B., Khoury, G., and Denhardt, D. (1975). Physical Map and Strand Polarity of Specific Fragments of Adenovirus-Associated Virus DNA Produced by Endonuclease R *EcoRI. J Virol *16*, 559–568.

Carter, B., Laughlin, C., de la Maza, L., and Myers, M. (1979). Adeno-associated virus auto-interference. Virology *92*, 449–462.

Carter, B., Trempe, J., and Mendelson, E. (1990). Adeno-associated virus gene expression and regulation.

Carty, S.M., Goldstrohm, a C., Suñé, C., Garcia-Blanco, M. a, and Greenleaf, a L. (2000). Protein-interaction modules that organize nuclear function: FF domains of CA150 bind the phosphoCTD of RNA polymerase II. Proc. Natl. Acad. Sci. U. S. A. *97*, 9015–9020.

Carvalho, T., Seeler, J.S., Öhman, K., Jordan, P., Pettersson, U., Akusjärvi, G., Carmo-Fonseca, M., and Dejean, A. (1995). Targeting of adenovirus E1A and E4-ORF3 proteins to nuclear matrix-associated PML bodies. J. Cell Biol. *131*, 45–56.

Casper, J.M., Timpe, J.M., Dignam, J.D., and Trempe, J.P. (2005). Identification of an adeno-associated virus Rep protein binding site in the adenovirus E2a promoter. J. Virol. *79*, 28–38.

Cataldi, M.P., and McCarty, D.M. (2010). Differential effects of DNA double-strand break repair pathways on single-strand and self-complementary adeno-associated virus vector genomes. J. Virol. *84*, 8673–8682.

Cataldi, M.P., and McCarty, D.M. (2013). Hairpin-end conformation of adeno-associated virus genome determines interactions with DNA-repair pathways. Gene Ther *20*, 686–693.

Cervelli, T., Palacios, J.A., Zentilin, L., Mano, M., Schwartz, R. a, Weitzman, M.D., and Giacca, M. (2008). Processing of recombinant AAV genomes occurs in specific nuclear structures that overlap with foci of DNA-damage-response proteins. J Cell Sci *121*, 349–357.

Ceulemans, H., and Bollen, M. (2004). Functional diversity of protein phosphatase-1, a cellular economizer and reset button. Physiol. Rev. *84*, 1–39.

Chandler, M., de la Cruz, F., Dyda, F., Hickman, A.B., Moncalian, G., and Ton-Hoang, B. (2013). Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. Nat. Rev. Microbiol. *11*, 525–538.

Chang, C.-J.C.-W.C.-J.C.-W., Chou, H.-Y., Lin, Y.-S., Huang, K.-H., Chang, C.-J.C.-W.C.-J.C.-W., Hsu, T.-C., and Lee, S.-C. (2008). Phosphorylation at Ser473 regulates heterochromatin protein 1 binding and corepressor function of TIF1beta/KAP1. BMC Mol. Biol. *9*, 61.

Chang, L.S., Shi, Y., and Shenk, T. (1989). Adeno-associated virus P5 promoter contains an adenovirus E1A-inducible element and a binding site for the major late transcription factor. J Virol *63*, 3479–3488.

Chang, P.C., Fitzgerald, L.D., Van Geelen, A., Izumiya, Y., Ellison, T.J., Wang, D.H., Ann, D.K., Luciw, P.A., and Kung, H.J. (2009). Kruppel-associated box domain-associated protein-1 as a latency regulator for Kaposi's sarcoma-associated herpesvirus and its modulation by the viral protein kinase. Cancer Res *69*, 5681–5689.

Chau, C.M., Zhang, X.Y., McMahon, S.B., and Lieberman, P.M. (2006). Regulation of Epstein-Barr virus latency type by the chromatin boundary factor CTCF. J Virol *80*, 5723–5732.

Chejanovsky, N., and Carter, B.J. (1989). Mutagenesis of an AUG codon in the adeno-associated virus rep gene: Effects on viral DNA replication. Virology *173*, 120–128.

Chejanovsky, N., and Carter, B.J. (1990). Mutation of a consensus purine nucleotide binding site in the adeno-associated virus rep gene generates a dominant negative phenotype for DNA replication. J Virol *64*, 1764–1770.

Chellappan, S., Kraus, V.B., Kroger, B., Munger, K., Howley, P.M., Phelps, W.C., and Nevins, J.R. (1992). Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product. Proc. Natl. Acad. Sci. U. S. A. *89*, 4549–4553.

Chen, C.-L., Jensen, R.L., Schnepp, B.C., Connell, M.J., Shell, R., Sferra, T.J., Bartlett, J.S., Clark, K.R., and Johnson, P.R. (2005). Molecular characterization of adeno-associated viruses infecting children. J. Virol. *79*, 14781–14792.

Chen, L., Cai, Y., Jin, G., Florens, L., Swanson, S.K., Washburn, M.P., Conaway, J.W., and Conaway, R.C. (2011). Subunit organization of the human INO80 chromatin remodeling complex: An evolutionarily conserved core complex catalyzes ATP-dependent nucleosome remodeling. J. Biol. Chem. *286*, 11283–11289.

Cheng, C.Y., Gilson, T., Wimmer, P., Schreiner, S., Ketner, G., Dobner, T., Branton, P.E., and Blanchette, P. (2013). Role of E1B55K in E4orf6/E1B55K E3 ligase complexes formed by different human adenovirus serotypes. J. Virol. *87*, 6232–6245.

Cheung, A.K., Hoggan, M.D., Hauswirth, W.W., and Berns, K.I. (1980). Integration of the adeno-associated virus genome into cellular DNA in latently infected human Detroit 6 cells. J. Virol. *33*, 739–748.

Chiorini, J.A., Yang, L., Safer, B., and Kotin, R.M. (1995). Determination of adeno-associated virus Rep68 and Rep78 binding sites by random sequence oligonucleotide selection. J. Virol. *69*, 7334–7338.

Chiorini, J.A., Zimmermann, B., Yang, L., Smith, R.H., Ahearn, A., Herberg, F., and Kotin, R.M. (1998). Inhibition of PrKX, a novel protein kinase, and the cyclic AMP- dependent protein kinase PKA by the regulatory proteins of adeno- associated virus type 2. Mol. Cell. Biol. *18*, 5921–5929.

Chiorini, J. a, Wiener, S.M., Owens, R. a, Kyöstió, S.R., Kotin, R.M., and Safer, B. (1994). Sequence requirements for stable binding and function of Rep68 on the adeno-associated virus type 2 inverted terminal repeats. J. Virol. *68*, 7448–7457.

Chirmule, N., Propert, K., Magosin, S., Qian, Y., Qian, R., and Wilson, J. (1999). Immune responses to adenovirus and adeno-associated virus in humans. Gene Ther. *6*, 1574–1583.
Choi, V.W., McCarty, D.M., and Samulski, R.J. (2006). Host cell DNA repair pathways in adeno-associated viral genome processing. J. Virol. *80*, 10346–10356.

Cohen, P.T.W. (2002). Protein phosphatase 1--targeted in many directions. J. Cell Sci. *115*, 241–256.

Coiras, M., Montes, M., Montanuy, I., Rosa López-Huertas, M., Mateos, E., Le Sommer, C., Garcia-Blanco, M.A., Hernández-Munain, C., Alcamí, J., and Suñé, C. (2013). Transcription elongation regulator 1 (TCERG1) regulates competent RNA polymerase II-mediated elongation of HIV-1 transcription and facilitates efficient viral replication. Retrovirology *10*, 1.

Collaco, R.F., Bevington, J.M., Bhrigu, V., Kalman-Maltese, V., and Trempe, J.P. (2009). Adeno-associated virus and adenovirus coinfection induces a cellular DNA damage and repair response via redundant phosphatidylinositol 3-like kinase pathways. Virology *392*, 24–33.

Conn, K.L., and Schang, L.M. (2013). Chromatin dynamics during lytic infection with herpes simplex virus 1. Viruses *5*, 1758–1786.

Costello, E., Saudan, P., Winocour, E., Pizer, L., and Beard, P. (1997). High mobility group chromosomal protein 1 binds to the adeno-associated virus replication protein (Rep) and promotes Rep-mediated site-specific cleavage of DNA, ATPase activity and transcriptional repression. EMBO J *16*, 5943–5954.

Cougot, D., Allemand, E., Riviere, L., Benhenda, S., Duroure, K., Levillayer, F., Muchardt, C., Buendia, M. -a. A., and Neuveut, C. (2012). Inhibition of PP1 phosphatase activity by HBx: a mechanism for the activation of hepatitis B virus transcription. Sci Signal *5*, ra1.

Crawford, L., Follett, E., Burdon, M., and McGeoch, D. (1969). The DNA of a Minute Virus of Mice. J Gen Virol *4*, 37–46.

Davis, M.D., Wu, J., and Owens, R. a (2000). Mutational analysis of adeno-associated virus type 2 Rep68 protein endonuclease activity on partially single-stranded substrates. J. Virol. *74*, 2936–2942.

Davis, M.E., Wang, M.K., Rennick, L.J., Full, F., Gableske, S., Mesman, A.W., Gringhuis, S.I., Geijtenbeek, T.B., Duprex, W.P., and Gack, M.U. (2014). Antagonism of the phosphatase PP1 by the measles virus V protein is required for innate immune escape of MDA5. Cell Host Microbe *16*, 19–30.

Dehal, P., Predki, P., Olsen, a S., Kobayashi, a, Folta, P., Lucas, S., Land, M., Terry, a, Ecale Zhou, C.L., Rash, S., et al. (2001). Human chromosome 19 and related regions in mouse: conservative and lineage-specific evolution. Science *293*, 104–111.

Denhardt, D.T., Eisenberg, S., Bartok, K., and Carter, B.J. (1976). Multiple structures of adeno-associated virus DNA: analysis of terminally labeled molecules with endonuclease R-Hae III. J. Virol. *18*, 672–684.

DeRijck, J., deKogel, C., Demeulemeester, J., Vets, S., ElAshkar, S., Malani, N., Bushman, F.D., Landuyt, B., Husson, S.J., Busschots, K., et al. (2013). The BET Family of Proteins Targets Moloney Murine Leukemia Virus Integration near Transcription Start Sites. Cell Rep. *5*, 886–894.

Diamant, G., and Dikstein, R. (2013). Transcriptional Control by NF-κB: Elongation in Focus. Biochim. Biophys. Acta - Gene Regul. Mech. *1829*, 937–945.

DiMattia, M. a, Nam, H.-J., Van Vliet, K., Mitchell, M., Bennett, A., Gurda, B.L., McKenna, R., Olson, N.H., Sinkovits, R.S., Potter, M., et al. (2012). Structural insight into the unique properties of adeno-associated virus serotype 9. J. Virol. *86*, 6947–6958.

Douar, A., Poulard, K., Stockholm, D., and Danos, O. (2001). Intracellular Trafficking of Adeno-Associated Virus Vectors : Routing to the Late Endosomal Compartment and Proteasome Degradation Intracellular Trafficking of Adeno-Associated Virus Vectors : Routing to the Late Endosomal Compartment and Proteasome Degra. Society *75*, 1824–1833.

Le Douarin, B., Nielsen, a L., Garnier, J.M., Ichinose, H., Jeanmougin, F., Losson, R., and Chambon, P. (1996). A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. EMBO J. *15*, 6701–6715.

Doucas, V., Ishov, A.M., Romo, A., Juguilon, H., Weitzman, M.D., Evans, R.M., and Maul, G.G. (1996). Adenovirus replication is coupled with the dynamic properties of the PML nuclear structure. Genes Dev. *10*, 196–207.

Drouin, L.M., and Agbandje-McKenna, M. (2013). Adeno-associated virus structural biology as a tool in vector development. Future Virol. *8*, 1183–1199.

Duan, D., Li, Q., Kao, A.W., Yue, Y., Pessin, J.E., and Engelhardt, J.F. (1999). Dynamin is required for recombinant adeno-associated virus type 2 infection. J. Virol. *73*, 10371–10376.

Dubielzig, R., King, J.A., Weger, S., Kern, A., and Kleinschmidt, J.A. (1999). Adeno-associated virus type 2 protein interactions: formation of pre-encapsidation complexes. J Virol *73*, 8989–8998.

Ducroux, A., Benhenda, S., Riviere, L., Semmes, O.J., Benkirane, M., Neuveut, C., Rivière, L., Semmes, O.J., Benkirane, M., and Neuveut, C. (2014). The Tudor domain protein Spindlin1 is involved in intrinsic antiviral defense against incoming hepatitis B Virus and herpes simplex virus type 1. PLoS Pathog *10*, e1004343.

Dupont, J.-B., Tournaire, B., Georger, C., Marolleau, B., Jeanson-Leh, L., Ledevin, M., Lindenbaum, P., Lecomte, E., Cogné, B., Dubreil, L., et al. (2015). Short-lived recombinant adeno-associated virus transgene expression in dystrophic muscle is associated with oxidative damage to transgene mRNA. Mol. Ther. Methods Clin. Dev. *2*, 15010.

Dutheil, N., and Linden, R.M. (2006). Site-specific integration by adeno-associated virus. In Parvoviruses, J.R. Kerr, S.F. Cotmore, M.E. Bloom, R.M. Linden, and C.R. Parrish, eds. (London, UK: Hodder Arnold), pp. 213–237.

Dutheil, N., Yoon-Robarts, M., Ward, P., Henckaerts, E., Skrabanek, L., Berns, K.I., Campagne, F., and Linden, R.M. (2004). Characterization of the Mouse Adeno-Associated Virus AAVS1 Ortholog. J. Virol. *78*, 8917–8921.

Dutheil, N., Henckaerts, E., Kohlbrenner, E., and Linden, R.M. (2009). Transcriptional analysis of the adeno-associated virus integration site. J. Virol. *83*, 12512–12525.

Dutheil, N., Smith, S.C., Agundez, L., Vincent-Mistiaen, Z.I., Escalante, C.R., Linden, R.M., and Henckaerts, E. (2014). Adeno-associated virus Rep represses the human integration site promoter by two pathways that are similar to those required for the regulation of the viral p5 promoter. J Virol *88*, 8227–8241.

Earley, L.F., Kawano, Y., Adachi, K., Sun, X.-X., Dai, M.-S., and Nakai, H. (2015). Identification and characterization of nuclear and nucleolar localization signals in the adeno-associated virus serotype 2 assembly-activating protein. J. Virol. *89*, 3038–3048.

Ecco, G., Cassano, M., Kauzlaric, A., Duc, J., Coluccio, A., Offner, S., Imbeault, M., Rowe, H.M., Turelli, P., and Trono, D. (2016). Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. Dev. Cell *36*, 611–623.
Emerson, R.O., and Thomas, J.H. (2009). Adaptive evolution in zinc finger transcription factors. PLoS Genet. *5*.

Enemark, E.J., Chen, G., Vaughn, D.E., Stenlund, A., and Joshua-Tor, L. (2000). Crystal structure of the DNA binding domain of the replication initiation protein E1 from papillomavirus. Mol Cell *6*, 149–158.

Evans, J., and Hearing, P. (2005). Relocalization of the Mre11-Rad50-Nbs1 complex by the adenovirus E4 ORF3 protein is required for viral replica- tion. J Virol *79*, 6207–6215.

Fahraeus, R., Palmqvist, L., Nerdstedt, A., Farzad, S., Rymo, L., and Lain, S. (1994). Response to cAMP levels of the Epstein-Barr virus EBNA2-inducible LMP1 oncogene and EBNA2 inhibition of a PP1-like activity. EMBO J *13*, 6041–6051.

Feuer, G., Taketo, M., Hanecak, R.C., and Fan, H. (1989). Two blocks in Moloney murine leukemia virus expression in undifferentiated F9 embryonal carcinoma cells as determined by transient expression assays. J. Virol. *63*, 2317–2324.

Fisher, K.J., Gao, G.P., Weitzman, M.D., DeMatteo, R., Burda, J.F., and Wilson, J.M. (1996). Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. J. Virol. *70*, 520–532.

Flamand, L., and Menezes, J. (1996). Cyclic AMP-responsive element-dependent activation of Epstein-Barr virus zebra promoter by human herpesvirus 6. J. Virol. *70*, 1784–1791.

Flanagan, J.R., Krieg, A.M., Max, E.E., and Khan, A.S. (1989). Negative control region at the 5' end of murine leukemia virus long terminal repeats. Mol. Cell. Biol. *9*, 739–746.

Flotte, T., Solow, R., Owens, R., Afione, S., Zeitlin, P., and Carter, B. (1992). Gene expression from adeno-associated virus vectors in airway epithelial cells. Am. J. Respir. Cell. Mol. Biol. *7*, 349–356.

Flotte, T., Zeitlin, P., Reynolds, T., Heald, A., Pederson, P., Beck, S., Conrad, C., Brass-Ernst, L., Humphries, M., Sullivan, K., et al. (2003). phase I trial of intranasal and endobronchial administration of a recombinant adeno-associated virus serotype 2 (rAAV2)-CFTR vector in adult cystic fibrosis pateients: a two part clinical study. Hum Gne Ther *14*, 1079–1088.

Flotte, T.R., Afione, S.A., Conrad, C., McGrath, S.A., Solow, R., Oka, H., Zeitlin, P.L., Guggino, W.B., and Carter, B.J. (1993). Stable in vivo expression of the cystic fibrosis transmembrane conductance regulator with an adeno-associated virus vector. Proc Natl Acad Sci U S A *90*, 10613–10617.

Forrester, N.A., Sedgwick, G.G., Thomas, A., Blackford, A.N., Speiseder, T., Dobner, T., Byrd, P.J., Stewart, G.S., Turnell, A.S., and Grand, R.J.A. (2011). Serotype-specific inactivation of the cellular DNA damage response during adenovirus infection. J. Virol. *85*, 2201–2211.

Forrester, N.A., Patel, R.N., Speiseder, T., Groitl, P., Sedgwick, G.G., Shimwell, N.J., Seed, R.I., Catnaigh, P.O., McCabe, C.J., Stewart, G.S., et al. (2012). Adenovirus E4orf3 targets transcriptional intermediary factor 1gamma for proteasome-dependent degradation during infection. J Virol *86*, 3167–3179.

Frauer, C., Rottach, A., Meilinger, D., Bultmann, S., Fellinger, K., Hasenöder, S., Wang, M., Qin, W., Söding, J., Spada, F., et al. (2011). Different binding properties and function of CXXC zinc finger domains in Dnmt1 and Tet1. PLoS One *6*.

Friedman, J.R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G., Rauscher 3rd, F.J., and Rauscher, F.J. (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev *10*, 2067–2078.

Friedman-Einat, M., Grossman, Z., Mileguir, F., Smetana, Z., Ashkenazi, M., Barkai, G., Varsano, N., Glick, E., and Mendelson, E. (1997). Detection of adeno-associated virus type 2 sequences in the human genital tract. J. Clin. Microbiol. *35*, 71–78.

Frietze, S., O'Geen, H., Blahnik, K.R., Jin, V.X., and Farnham, P.J. (2010). ZNF274 recruits the histone methyltransferase SETDB1 to the 39 ends of ZNF genes. PLoS One *5*.

Garvin, A.J., Densham, R.M., Blair-Reid, S.A., Pratt, K.M., Stone, H.R., Weekes, D., Lawrence, K.J., and Morris, J.R. (2013). The deSUMOylase SENP7 promotes chromatin relaxation for homologous recombination DNA repair. EMBO Rep. *14*, 975–983.

Ge, H., and Roeder, R.G. (1994). Purification, cloning, and characterization of a human coactivator, PC4, that mediates transcriptional activation of class II genes. Cell *78*, 513–523.

Geoffroy, Epstein, Toublanc, and Moullier (2004a). Herpes simplex virus type 1 ICP0 protein mediates activation of adeno-associated virus type 2 rep …. J. Virol.

Geoffroy, M.C., Epstein, A.L., Toublanc, E., Moullier, P., and Salvetti, A. (2004b). Herpes simplex virus type 1 ICP0 protein mediates activation of adeno-associated virus type 2 rep gene expression from a latent integrated form. J Virol *78*, 10977–10986.

Georg-Fries, B., Biederlack, S., Wolf, J., and Zur Hausen, H. (1984). Analysis of proteins, helper dependence, and seroepidemiology of a new human parvovirus. Virology *134*, 64–71.

Giberson, A.N., Davidson, A.R., and Parks, R.J. (2012). Chromatin structure of adenovirus DNA throughout infection. Nucleic Acids Res. *40*, 2369–2376.

Giebler, H.A., Lemasson, I., and Nyborg, J.K. (2000). p53 recruitment of CREB binding protein mediated through phosphorylated CREB: a novel pathway of tumor suppressor regulation. Mol. Cell. Biol. *20*, 4849–4858.

Gilchrist, C., Orten, D., and Hinrichs, S. (1996). Evidence for the Role of Cyclic AMP-Responsive Elements in Human Virus Replication and Disease. J Biomed Sci *3*, 293–306.

Gill, G. (2005). Something about SUMO inhibits transcription. Curr. Opin. Genet. Dev. *15*, 536–541.

Giraud, C., Winocour, E., and Berns, K.I. (1994). Site-specific integration by adeno-associated virus is directed by a cellular DNA sequence. Proc. Natl. Acad. Sci. U. S. A. *91*, 10039–10043.

Girod, A., Wobus, C., Zadori, Z., Ried, M., Like, K., Tijssen, P., Kleinschmidt, J., and Hallek, M. (2002). The VP1 capsid protein of adeno-associated virus type 2 is carrying a phospholipase A2 domain required for virus infectivity. J Gen Virol *83*, 973–978.

Glauser, D.L., Strasser, R., Laimbacher, A.S., Saydam, O., Clément, N., Linden, R.M., Ackermann, M., and Fraefel, C. (2007). Live covisualization of competing adeno-associated virus and herpes simplex virus type 1 DNA replication: molecular mechanisms of interaction. J. Virol. *81*, 4732–4743.

Glauser, D.L., Seyffert, M., Strasser, R., Franchini, M., Laimbacher, A.S., Dresch, C., de Oliveira, A.P., Vogel, R., Büning, H., Salvetti, A., et al. (2010). Inhibition of herpes simplex virus type 1 replication by adeno-associated virus rep proteins depends on their combined DNA-binding and ATPase/helicase activities. J Virol *84*, 3808–3824.

Goldstrohm, A.C., Albrecht, T.R., Sune, C., Bedford, M.T., and Garcia-Blanco, M.A. (2001). The Transcription Elongation Factor CA150 Interacts with RNA Polymerase II and the Pre-mRNA Splicing Factor SF1. Mol. Cell. Biol. *21*, 7617–7628.

Goodarzi, A.A., Noon, A.T., Deckbar, D., Ziv, Y., Shiloh, Y., Lubrich, M., and Jeggo, P.A. (2008). ATM Signaling Facilitates Repair of DNA Double-Strand Breaks Associated with Heterochromatin. Mol. Cell *31*, 167–177.

Goodarzi, A.A., Kurka, T., and Jeggo, P.A. (2011). KAP-1 phosphorylation regulates CHD3 nucleosome remodeling during the DNA double-strand break response. Nat Struct Mol Biol *18*, 831–839.

Green, M.R., and Roeder, R.G. (1980). Definition of a novel promoter for the major adenovirus-associated virus mRNA. Cell *22*, 231–242.

Grieger, J.C., Snowdy, S., and Samulski, R.J. (2006). Separate basic region motifs within the adeno-associated virus capsid proteins are essential for infectivity and assembly. J. Virol. *80*, 5199–5210.

Grimm, D., Kern, A., Rittner, K., and Kleinschmidt, J.A. (1998). Novel tools for production and purification of recombinant adenoassociated virus vectors. Hum. Gene Ther. *9*, 2745–2760.

Groner, A.C., Meylan, S., Ciuffi, A., Zangger, N., Ambrosini, G., Denervaud, N., Bucher, P., Trono, D., Dénervaud, N., Bucher, P., et al. (2010). KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. PLoS Genet *6*, e1000869.

Grossman, Z., Mendelson, E., Brok-Simoni, F., Mileguir, F., Leitner, Y., Rechavi, G., and Ramot, B. (1992). Detection of adeno-associated virus type 2 in human peripheral blood cells. J. Gen. Virol. *73*, 961–966.

Guo, B., Stein, J.L., Van Wijnen, A.J., and Stein, G.S. (1997). ATF1 and CREB trans-activate a cell cycle regulated histone H4 gene at a distal nuclear matrix associated promoter element. Biochemistry *36*, 14447–14455.

Gupta, a, Jha, S., Engel, D. a, Ornelles, D. a, and Dutta, A. (2013). Tip60 degradation by adenovirus relieves transcriptional repression of viral transcriptional activator EIA. Oncogene *32*, 5017–5025.

Hagiwara, M., Alberts, A., Brindle, P., Meinkoth, J., Feramisco, J., Deng, T., Karin, M., Shenolikar, S., and Montminy, M. (1992). Transcriptional attenuation following cAMP induction requires PP-1-mediated dephosphorylation of CREB. Cell *70*, 105–113.

Hamilton, H., Gomos, J., Berns, K.I., and Falck-Pedersen, E. (2004). Adeno-associated virus site-specific integration and AAVS1 disruption. J. Virol. *78*, 7874–7882.

Han, L., Parmley, T.H., Keith, S., Kozlowski, K.J., Smith, L.J., and Hermonat, P.L. (1996). High prevalence of adeno-associated virus (AAV) type 2 rep DNA in cervical materials: AAV may be sexually transmitted. Virus Genes *12*, 47–52.

Handa, H., Shiroki, K., and Shimojo, H. (1977). establishment and characterization of KB cell lines latently infected with adeno-associated virus type 1. Virology *82*, 84–92.

Harada, J.N., Shevchenko, A., Shevchenko, A., Pallas, D.C., and Berk, A.J. (2002). Analysis of the adenovirus E1B-55K-anchored proteome reveals its link to ubiquitination machinery. J. Virol. *76*, 9194–9206.

Hardwicke, M.A., and Sandri-Goldin, R. M. (1994) The herpes simplex virus regulatory protein ICP27 contributes to the decrease in cellular mRA levels during infection. J. Virol. *68*, 4797-4810.

Hargreaves, D.C., Horng, T., and Medzhitov, R. (2009). Control of Inducible Gene Expression by Signal-Dependent Transcriptional Elongation. Cell *138*, 129–145.

Haruki, H., Okuwaki, M., Miyagishi, M., Taira, K., and Nagata, K. (2006). Involvement of template-activating factor I/SET in transcription of adenovirus early genes as a positive-acting factor. J Virol *80*, 794–801.

Hauswirth, W., and Berns, K. (1978). Origin and Termination of adeno-associated virus DNA replication. Virology1977.

He, N., Jahchan, N.S., Hong, E., Li, Q., Bayfield, M.A., Maraia, R.J., Luo, K., and Zhou, Q. (2008). A La-Related Protein Modulates 7SK snRNP Integrity to Suppress P-TEFb-Dependent Transcriptional Elongation and Tumorigenesis. Mol. Cell *29*, 588–599.

Henckaerts, E., and Linden, R.M. (2010). Adeno-associated virus: a key to the human genome? Futur. Virol *5*, 555–574.

Henckaerts, E., Dutheil, N., Zeltner, N., Kattman, S., Kohlbrenner, E., Ward, P., Clément, N., Rebollo, P., Kennedy, M., Keller, G.M., et al. (2009). Site-specific integration of adeno-associated virus involves partial duplication of the target locus. Proc Natl Acad Sci U S A *106*, 7571–7576.

Hendrickx, A., Beullens, M., Ceulemans, H., Abt, T., Van Eynde, A., Nicolaescu, E., Lesage, B., and Bollen, M. (2009). Docking motif-guided mapping of the interactome of protein phosphatase-1. Chem Biol *16*, 365–371.

Hermanns, J., Schulze, A., Jansen-DBurr, P., Kleinshmidt, J., Schmidt, R., and zur Hausen, H. (1997). Infection of primary cells by adeno-associated virus type 2 results in a modulation of cell cycle-regulating proteins. J Virol *71*, 6020–6027.

Hermonat, P. (1994). Down-regulation of the human c-fos and c-myc proto-oncogene promoters by adeno-associated virus Rep78. Cancer Lett *81*, 129–136.

Hermonat, P.L. (1991). Inhibition of H-ras expression by the adeno-associated virus Rep78 transformation suppressor gene product. Cancer Res. *51*, 3373–3377.

Hermonat, P.L., Labow, M.A., Wright, R., Berns, K.I., and Muzyczka, N. (1984). Genetics of adeno-associated virus: isolation and preliminary characterization of adeno-associated virus type 2 mutants. J. Virol. *51*, 329–339.

Hermonat, P.L., Santin, A.D., and Batchu, R.B. (1996). The adeno-associated virus Rep78 major regulatory/transformation suppressor protein binds cellular Sp1 in vitro and evidence of a biological effect. Cancer Res *56*, 5299–5304.

Hernonat, P., Santin, A.D., Batchu, R.B., Zhan, D., Hermonat, P.L., Santin, A.D., Batchu, R.B., Zhan, D., Hermonat, P.L., Santin, A.D., et al. (1998). The adeno-associated virus Rep78 major regulatory protein binds the cellular TATA-binding protein in vitro and in vivo. Virology *245*, 120–127.

Hickman, A.B., Ronning, D.R., Kotin, R.M., and Dyda, F. (2002). Structural unity among viral origin binding proteins: Crystal structure of the nuclease domain of adeno-associated virus Rep. Mol. Cell *10*, 327–337.

Le Hir, M., Goyenvalle, A., Peccate, C., Précigout, G., Davies, K.E., Voit, T., Garcia, L., and Lorain, S. (2013). AAV genome loss from dystrophic mouse muscles during AAV-U7 snRNA-mediated exon-skipping therapy. Mol. Ther. *21*, 1551–1558.

Hoggan, M., Blacklow, N., and Rowe, W. (1966). Studies of small DNA viruses found in various adenovirus preparations: physical, biological, and immunological characteristics. Proc Natl Acad Sci U S A *55*, 1467–1472.

Hoggan, M., Thomas, G., and Johnson, F. (1972). Continuous carriage of adenovirus-associated virus genome in cell culture in the absence of helper adenovirus. Proceeding Fourth Lepetit Colloq. 41–47.

Hölscher, C., Kleinschmidt, J. a, and Bürkle, A. (1995). High-level expression of adeno-associated virus (AAV) Rep78 or Rep68 protein is sufficient for infectious-particle formation by a rep-negative AAV mutant. J. Virol. *69*, 6880–6885.

Hörer, M., Weger, S., Butz, K., Hoppe-Seyler, F., Geisen, C., Kleinschmidt, J. a, Horer, M., Weger, S., Butz, K., Hoppe-Seyler, F., et al. (1995). Mutational analysis of adeno-associated virus Rep protein-mediated inhibition of heterologous and homologous promoters. J. Virol. *69*, 5485–5496.

Hu, C., Zhang, S., Gao, X., Gao, X., Xu, X., Lv, Y., Zhang, Y., Zhu, Z., Zhang, C., Li, Q., et al. (2012). Roles of Kruppel-associated Box (KRAB)-associated Co-repressor KAP1 Ser-473 Phosphorylation in DNA Damage Response. J Biol Chem *287*, 18937–18952.

Hunter, L. a, and Samulski, R.J. (1992). Colocalization of adeno-associated virus Rep and capsid proteins in the nuclei of infected cells. J. Virol. *66*, 317–324.

Huser, D., Weger, S., and Heilbronn, R. (2002). Kinetics and frequency of adeno-associated virus site-specific integration into human chromosome 19 monitored by quantitative real-time PCR. J Virol *76*, 7554–7559.

Hüser, D., Gogol-Döring, A., Chen, W., and Heilbronn, R. (2014). AAV Type 2 Wild-Type and Vector-Mediated Genomic Integration Profiles in Human Diploid Fibroblasts Analyzed by 3rd Generation PacBio DNA Sequencing. J. Virol. *88*, JVI.01356-14-.

Im, D.S., and Muzyczka, N. (1990a). The AAV origin binding protein Rep68 is an ATP-dependent site-specific endonuclease with DNA helicase activity. Cell *61*, 447–457.

Im, D.S., and Muzyczka, N. (1990b). The AAV origin Bonding protein Rep 68 is an ATP-dependent site-specific endonuclease with DNA helicase activity. Cell *61*, 447–457.

Im, D.S., and Muzyczka, N. (1992). Partial purification of adeno-associated virus Rep78, Rep52, and Rep40 and their biochemical characterization. J. Virol. *66*, 1119–1128.

Ishizuka, M., Kawate, H., Takayanagi, R., Ohshima, H., Tao, R.H., and Hagiwara, H. (2005). A zinc finger protein TZF is a novel corepressor of androgen receptor. Biochem. Biophys. Res. Commun. *331*, 1025–1031.

Isogai, S., and Shirakawa, M. (2007). Protein modification by SUMO. Seikagaku *79*, 1120–1130.

Itokawa, Y., Yanagawa, T., Yamakawa, H., Watanabe, N., Koga, H., and Nagase, T. (2009). KAP1-independent transcriptional repression of SCAN-KRAB-containing zinc finger proteins. Biochem. Biophys. Res. Commun. *388*, 689–694.

Ivanov, A. V., Peng, H., Yurchenko, V., Yap, K.L., Negorev, D.G., Schultz, D.C., Psulkowski, E., Fredericks, W.J., White, D.E., Maul, G.G., et al. (2007). PHD Domain-Mediated E3 Ligase Activity Directs Intramolecular Sumoylation of an Adjacent Bromodomain Required for Gene Silencing. Mol. Cell *28*, 823–837.

Iyengar, S., and Farnham, P.J. (2011). KAP1 protein: an enigmatic master regulator of the genome. J Biol Chem *286*, 26267–26276.

Iyengar, S., Ivanov, A. V., Jin, V.X., Rauscher 3rd, F.J., Farnham, P.J., Rauscher, F.J., Farnham, P.J., Rauscher 3rd, F.J., and Farnham, P.J. (2011). Functional analysis of KAP1 genomic recruitment. Mol Cell Biol *31*, 1833–1847.

Jagiello, I., Beullens, M., Stalmans, W., and Bollen, M. (1995). Subunit structure and regulation of protein phosphatase-1 in rat liver nuclei. J. Biol. Chem. *270*, 17257–17263.

Jagiello, I., Beullens, M., Vulsteke, V., Wera, S., Sohlberg, B., Stalmans, W., von Gabain, a, and Bollen, M. (1997). NIPP-1, a nuclear inhibitory subunit of protein phosphatase-1, has RNA-binding properties. J. Biol. Chem. *272*, 22067–22071.

Jakobsson, J., Cordero, M.I., Bisaz, R., Groner, A.C., Busskamp, V., Bensadoun, J.C., Cammas, F., Losson, R., Mansuy, I.M., Sandi, C., et al. (2008). KAP1-Mediated Epigenetic Repression in the Forebrain Modulates Behavioral Vulnerability to Stress. Neuron *60*, 818–831.

James, J.A., Escalante, C.R., Yoon-Robarts, M., Edwards, T.A., Linden, R.M., and Aggarwal, A.K. (2003). Crystal structure of the SF3 helicase from adeno-associated virus type 2. Structure *11*, 1025–1035.

Janovitz, T., Klein, I.A., Oliveira, T., Mukherjee, P., Nussenzweig, M.C., Sadelain, M., and Falck-Pedersen, E. (2013). High-throughput sequencing reveals principles of adeno-associated virus serotype 2 integration. J. Virol. *87*, 8559–8568.

Jay, F.T., Laughlin, C.A., and Carter, B.J. (1981). Eukaryotic translational control: adeno-associated virus protein synthesis is affected by a mutation in the adenovirus DNA-binding protein. Proc. Natl. Acad. Sci. U. S. A. *78*, 2927–2931.

Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Tharien, C., Bergeron, D., Bourassa, S., Greenblatt, J., et al. (2007). Systematic Analysis of the Protein Interaction Network for the Human Transcription Machinery Reveals the Identity of the 7SK Capping Enzyme. Mol. Cell *27*, 262–274.

Jha, S., Shibata, E., and Dutta, A. (2008). Human Rvb1/Tip49 is required for the histone acetyltransferase activity of Tip60/NuA4 and for the downregulation of phosphorylation on H2AX after DNA damage. Mol Cell Biol *28*.

Jin, J., Cai, Y., Yao, T., Gottschalk, A.J., Florens, L., Swanson, S.K., Gutierrez, J.L., Coleman, M.K., Workman, J.L., Mushegian, A., et al. (2005). A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex. J. Biol. Chem. *280*, 41207–41212.

Jing, X.J., Kalman-Maltese, V., Cao, X., Yang, Q., and Trempe, J.P. (2001). Inhibition of adenovirus cytotoxicity, replication, and E2a gene expression by adeno-associated virus. Virology *291*, 140–151.

Joazeiro, C., and Weissman, A. (2000). RING finger proteins: mediators of ubiquitin ligase activity. Cell *102*, 549–552.

Johnson, J.S., and Samulski, R.J. (2009). Enhancement of adeno-associated virus infection by mobilizing capsids into and out of the nucleolus. J. Virol. *83*, 2632–2644.

Johnson, F., Blacklow, N., and Hoggan, M. (1972). Immunological Reactivity of Antisera Prepared Against the Sodium Dodecyl Sulfate-Treated Structural Polypeptides of Adenovirus- Associated Virus. J Virol *9*, 1017–1026.

Johnson, J.S., Li, C., DiPrimio, N., Weinberg, M.S., McCown, T.J., and Samulski, R.J. (2010). Mutagenesis of adeno-associated virus type 2 capsid protein VP1 uncovers new roles for basic amino acids in trafficking and cell-specific transduction. J. Virol. *84*, 8888–8902.

Jurvansuu, J., Raj, K., Stasiak, A., and Beard, P. (2005). Viral transport of DNA damage that mimics a stalled replication fork. J. Virol. *79*, 569–580.

Kailasan, S., Halder, S., Gurda, B., Bladek, H., Chipman, P.R., McKenna, R., Brown, K., and Agbandje-McKenna, M. (2015). Structure of an Enteric Pathogen, Bovine Parvovirus. J. Virol. *89*, 2603–2614.

Kaludov, N., Brown, K.E., Walters, R.W., Zabner, J., and Chiorini, J.A. (2001). Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. J. Virol. *75*, 6884–6893.

Kazemi, S., Papadopoulou, S., Li, S., Su, Q., Wang, S., Yoshimura, A., Matlashewski, G., Dever, T., and Koromilas, A. (2004). Control of alpha subunit of eukaryotic translation initiation factor 2 (eIF2 alpha) phosphorylation by the human papillomavirus type 18 E6 oncoprotein: implications for eIF2 alpha-dependent gene expression and cell death. Mol Cell Biol *24*, 3415–3429.

Kearns, W.G., Afione, S.A., Fulmer, S.B., Pang, M.C., Erikson, D., Egan, M., Landrum, M.J., Flotte, T.R., and Cutting, G.R. (1996). Recombinant adeno-associated virus (AAV-CFTR) vectors do not integrate in a site-specific fashion in an immortalized epithelial cell line. Gene Ther. *3*, 748–755.

Kern, A., Schmidt, K., Leder, C., Muller, O.J., Wobus, C.E., Bettinger, K., Von der Lieth, C.W., King, J.A., and Kleinschmidt, J.A. (2003). Identification of a heparin-binding motif on adeno-associated virus type 2 capsids. J Virol *77*, 11072–11081.

Kerur, N., Veettil, M.V., Sharma-Walia, N., Bottero, V., Sadagopan, S., Otageri, P., and Chandran, B. (2011). IFI16 acts as a nuclear pathogen sensor to induce the inflammasome in response to Kaposi Sarcoma-associated herpesvirus infection. Cell Host Microbe *9*, 363–375.

Kilham, L., and Olivier, L. (1959). A latent virus of rats isolated in tissue culture. Virology *7*, 428–437.

Kim, S., Yamamoto, J., Chen, Y., Aida, M., Wada, T., Handa, H., and Yamaguchi, Y. (2010). Evidence that cleavage factor Im is a heterotetrameric protein complex controlling alternative polyadenylation. Genes Cells *15*, 1003–1013.

Kim, S.S., Chen, Y.M., O'Leary, E., Witzgall, R., Vidal, M., and Bonventre, J. V (1996). A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins. Proc. Natl. Acad. Sci. U. S. A. *93*, 15299–15304.

Kim, Y.K., Bourgeois, C.F., Isel, C., Churcher, M.J., and Karn, J. (2002). Phosphorylation of the RNA polymerase II carboxyl-terminal domain by CDK9 is directly responsible for human immunodeficiency virus type 1 Tat-activated transcriptional elongation. Mol. Cell. Biol. *22*, 4622–4637.

King, J.A., Dubielzig, R., Grimm, D., and Kleinschmidt, J.A. (2001). DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. EMBO J *20*, 3282–3291.

Knipe, D.M., and Cliffe, A. (2008). Chromatin control of herpes simplex virus lytic and latent infection. Nat. Rev. Microbiol. *6*, 211–221.

Knipe, D.M., Lieberman, P.M., Jung, J.U., McBride, A.A., Morris, K. V, Ott, M., Margolis, D., Nieto, A., Nevels, M., Parks, R.J., et al. (2013). Snapshots: chromatin control of viral infection. Virology *435*, 141–156.

Koczot, F.J., Carter, B.J., Garon, C.F., and Rose, J. a (1973). Self-complementarity of

terminal sequences within plus or minus strands of adenovirus-associated virus DNA. Proc. Natl. Acad. Sci. U. S. A. *70*, 215–219.

Koonin, E. V., and Ilyina, T. V. (1992). Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA replication initiator proteins. J. Gen. Virol. *73*, 2763–2766.

Koonin, E. V., and Ilyina, T. V. (1993). Computer-assisted dissection of rolling circle DNA replication. BioSystems *30*, 241–268.

Kotin, R.M., and Berns, K.I. (1989). Organization of adeno-associated virus DNA in latently infected detroit 6 cells. Virology *170*, 460–467.

Kotin, R.M., Siniscalco, M., Samulski, R.J., Zhu, X.D., Hunter, L., Laughlin, C.A., McLaughlin, S., Muzyczka, N., Rocchi, M., and Berns, K.I. (1990). Site-specific integration by adeno-associated virus. Proc. Natl. Acad. Sci. U. S. A. *87*, 2211–2215.

Kotin, R.M., Menninger, J.C., Ward, D.C., and Berns, K.I. (1991). Mapping and direct visualization of a region-specific viral DNA integration site on chromosome 19q13-qter. Genomics *10*, 831–834.

Kotin, R.M., Linden, R.M., and Berns, K.I. (1992). Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. EMBO J. *11*, 5071–5078.

Kretzschmar, M., Kaiser, K., Lottspeich, F., and Meisterernst, M. (1994). A novel mediator of class II gene transcription with homology to viral immediate-early transcriptional regulators. Cell *78*, 525–534.

Kronenberg, S., Böttcher, B., von der Lieth, C.W., Bleker, S., and Kleinschmidt, J.A. (2005). A conformational change in the adeno-associated virus type 2 capsid leads to the exposure of hidden VP1 N termini. J. Virol. *79*, 5296–5303.

Krueger, B.J., Jeronimo, C., Roy, B.B., Bouchard, A., Barrandon, C., Byers, S.A., Searcey, C.E., Cooper, J.J., Bensaude, O., Cohen, É.A., et al. (2008). LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. Nucleic Acids Res. *36*, 2219–2229.

Kube, D.M., Ponnazhagan, S., and Srivastava, a (1997). Encapsidation of adeno-associated virus type 2 Rep proteins in wild-type and recombinant progeny virions: Rep-mediated growth inhibition of primary human cells. J. Virol. *71*, 7361–7371.

Kuo, C.Y., Li, X., Kong, X., Luo, C., Chang, C., Chung, Y., Shih, H., Li, K., and Ann, D. (2014). An Arginine-rich Motif of Ring Finger Protein 4 (RNF4) Oversees the Recruitment and Degradation of the Phosphorylated and SUMOylated Krüppel-associated Box Domain-associated Protein 1 (KAP1)/TRIM28 Protein during Genotoxic Stress. J Biol Chem *289*, 20757–20772.

Kwok, R.P., Laurance, M.E., Lundblad, J.R., Goldman, P.S., Shih, H., Connor, L.M., Marriott, S.J., and Goodman, R.H. (1996). Control of cAMP-regulated enhancers by the viral transactivator Tax through CREB and the co-activator CBP. Nature *380*, 642–646.

Kwong, A.D. and Frenkel, N. (1987). Herpes simplex virus-infected cells contain a function(s) that destabilizes both host and viral mRNAs. Proc. Natl. Acad. Sci. U. S. A. *84*,1926-1930.

Kyostio, S.R., Wonderling, R.S., and Owens, R.A. (1995). Negative regulation of the adeno-associated virus (AAV) P5 promoter involves both the P5 rep binding site and the consensus ATP-binding motif of the AAV Rep68 protein. J Virol *69*, 6787–6796.

Kyöstiö, S.R., Owens, R.A., Weitzman, M.D., Antoni, B.A., Chejanovsky, N., and Carter, B.J. (1994). Analysis of adeno-associated virus (AAV) wild-type and mutant Rep proteins for their abilities to negatively regulate AAV p5 and p19 mRNA levels. J. Virol. *68*, 2947–2957.

Labow, M.A., Hermonat, P.L., and Berns, K.I. (1986). Positive and negative autoregulation of the adeno-associated virus type 2 genome. J. Virol. *60*, 251–258.

Lacasse, J.J., and Schang, L.M. (2012). Herpes simplex virus 1 DNA is in unstable nucleosomes throughout the lytic infection cycle, and the instability of the nucleosomes is independent of DNA replication. J. Virol. *86*, 11287–11300.

Lackner, D.F., and Muzyczka, N. (2002). Studies of the mechanism of transactivation of the adeno-associated virus p19 promoter by Rep protein. J. Virol. *76*, 8225–8235.

Lallemand-Breitenbach, V., and de Thé, H. (2010). PML nuclear bodies. Cold Spring Harb. Perspect. Biol. *2*.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Laughlin, C.A., Myers, M.W., Risin, D.L., and Carter, B.J. (1979a). Defective-interfering particles of the human parvovirus adeno-associated virus. Virology *94*, 162–174.

Laughlin, C.A., Tratschin, J.D., Coon, H., and Carter, B.J. (1983). Cloning of infectious adeno-associated virus genomes in bacterial plasmids. Gene *23*, 65–73.

Laughlin, C.A., Cardellichio, C.B., and Coon, H.C. (1986). Latent infection of KB cells with adeno-associated virus type 2. J. Virol. *60*, 515–524.

Laughlin, C. a, Westphal, H., and Carter, B.J. (1979b). Spliced adenovirus-associated virus RNA. Proc. Natl. Acad. Sci. U. S. A. *76*, 5567–5571.

Lechner, M.S., Begg, G.E., Speicher, D.W., and Rauscher, F.J. (2000). Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: direct chromoshadow domain-KAP-1 corepressor interaction is essential. Mol. Cell. Biol. *20*, 6449–6465.

Lee, D.-H., Goodarzi, A.A., Adelmant, G.O., Pan, Y., Jeggo, P.A., Marto, J.A., and Chowdhury, D. (2012). Phosphoproteomic analysis reveals that PP4 dephosphorylates KAP-1 impacting the DNA damage response. EMBO J. *31*, 2403–2415.

Lee, J.H., You, J., Dobrota, E., and Skalnik, D.G. (2010). Identification and characterization of a novel human PP1 phosphatase complex. J. Biol. Chem. *285*, 24466–24476.

Lee, J.S., Galvin, K.M., See, R.H., Eckner, R., Livingston, D., Moran, E., and Shi, Y. (1995). Relief of YY1 transcriptional repression by adenovirus E1A is mediated by E1A-associated protein p300. Genes Dev. *9*, 1188–1198.

Lee, Y.K., Thomas, S.N., Yang, A.J., and Ann, D.K. (2007). Doxorubicin down-regulates Krüppel-associated box domain-associated protein 1 sumoylation that relieves its transcription repression on p21 WAF1/CIP1 in Breast cancer MCF-7 cells. J. Biol. Chem. *282*, 1595–1606.

van Leeuwen, H., Okuwaki, M., Hong, R., Chakravarti, D., Nagata, K., and O'Hare, P. (2003). Herpes simplex virus type 1 tegument protein VP22 interacts with TAF-I proteins and inhibits nucleosome assembly but not regulation of histone acetylation by INHAT. J Gen Virol *84*, 2501–2510.

Lefebvre, R.B., Riva, S., and Berns, K.I. (1984). Conformation takes precedence over sequence in adeno-associated virus DNA replication. *4*, 1416–1419.

Léger, A., Le Guiner, C., Nickerson, M.L., McGee Im, K., Ferry, N., Moullier, P., Snyder, R.O., and Penaud-Budloo, M. (2011). Adeno-associated viral vector-mediated transgene expression is independent of DNA methylation in primate liver and skeletal muscle. PLoS One *6*, e20881.

Leib, D., Nadeau, K., Rundle, S., Shaffer, P., and Schaffer, P. (1991). The promoter of the latency-associated transcripts of herpes simplex virus type 1 contains a functional cAMP-response element: role of the latency-associated transcripts and cAMP in reactivation of viral latency. Proc Natl Acad Sci U S A *88*, 48–52.

Lentz, T.B., Samulski, R.J., and Samulski, R.J. (2015). Insight into the mechanism of inhibition of adeno-associated virus by the Mre11/Rad50/Nbs1 complex. J. Virol. *89*, 181–194.

Lerch, T.F., Xie, Q., and Chapman, M.S. (2010). The structure of adeno-associated virus serotype 3B (AAV-3B): Insights into receptor binding and immune evasion. Virology *403*, 26–36.

Levy, H.C., Bowman, V.D., Govindasamy, L., McKenna, R., Nash, K., Warrington, K., Chen, W., Muzyczka, N., Yan, X., Baker, T.S., et al. (2009). Heparin binding induces conformational changes in Adeno-associated virus serotype 2. J. Struct. Biol. *165*, 146–156.

Lewis, B. a, Tullis, G., Seto, E., Horikoshi, N., Weinmann, R., and Shenk, T. (1995). Adenovirus E1A proteins interact with the cellular YY1 transcription factor. J. Virol. *69*, 1628–1636.

Leza, M.A., and Hearing, P. (1989). Independent cyclic AMP and E1A induction of adenovirus early region 4 expression. J Virol *63*, 3057–3064.

Li, B., Carey, M., and Workman, J.L. (2007a). The Role of Chromatin during Transcription. Cell *128*, 707–719.

Li, H., Rauch, T., Chen, Z.X., Szabó, P.E., Riggs, A.D., and Pfeifer, G.P. (2006). The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. J. Biol. Chem. *281*, 19489–19500.

Li, T., Diner, B. a., Chen, J., and Cristea, I.M. (2012). Acetylation modulates cellular distribution and DNA sensing ability of interferon-inducible protein IFI16. Proc. Natl. Acad. Sci. *109*, 10558–10563.

Li, T., Chen, J., and Cristea, I.M. (2013). Human cytomegalovirus tegument protein pUL83 inhibits IFI16-mediated DNA sensing for immune evasion. Cell Host Microbe *14*, 591–599.

Li, X., Lee, Y.K., Jeng, J.C., Yen, Y., Schultz, D.C., Shih, H.M., and Ann, D.K. (2007b). Role for KAP1 serine 824 phosphorylation and sumoylation/desumoylation switch in regulating KAP1-mediated transcriptional repression. J Biol Chem *282*, 36177–36189.

Li, X., Lin, H.H., Chen, H., Xu, X., Shih, H.M., and Ann, D.K. (2010). SUMOylation of the transcriptional co-repressor KAP1 is regulated by the serine and threonine phosphatase PP1. Sci Signal *3*, ra32.

Li, Y., Zhang, C., Chen, X., Yu, J., Wang, Y., Yang, Y., Du, M., Jin, H., Ma, Y., He, B., et al. (2011). ICP34.5 protein of herpes simplex virus facilitates the initiation of protein translation by bridging eukaryotic initiation factor 2alpha (eIF2alpha) and protein phosphatase 1. J Biol Chem *286*, 24785–24792.

Li, Z., Wang, D., Na, X., Schoen, S.R., Messing, E.M., and Wu, G. (2003). The VHL protein recruits a novel KRAB-A domain protein to repress HIF-1α transcriptional activity. EMBO J. *22*, 1857–1867.

Lieberman, P. (2016). Epigenetics and Genetics of Viral Latency. Cell Host Microbe *19*, 619–628.

Lieberman, P.M. (2008). Chromatin organization and virus gene expression. J. Cell. Physiol. *216*, 295–302.

Linden, R., Ward, P., Winocour, E., and Berns, K. (1996a). Site specific integration by adeno-associated virus. Proc Natl Acad Sci U S A *93*, 11288–11294.

Linden, R.M., Winocour, E., and Berns, K.I. (1996b). The recombination signals for adeno-associated virus site-specific integration. Proc. Natl. Acad. Sci. U. S. A. *93*, 7966–7972.

Loh, T.P., Sievert, L.L., and Scott, R.W. (1987). Proviral sequences that restrict retroviral expression in mouse embryonal carcinoma cells. Mol. Cell. Biol. *7*, 3775–3784.

Loh, T.P., Sievert, L.L., and Scott, R.W. (1988). Negative regulation of retrovirus expression in embryonal carcinoma cells mediated by an intragenic domain. J. Virol. *62*, 4086–4095.

Lombardo, E., Ramirez, J., Agbandje-McKenna, M., and Almendral, J. (2000). a beta-strand motif drives capsid protein oligomers of the parvovrirus minute virus of mice into the nucleus for viral assembly. J Virol *74*, 3804–3814.

Lubeck, M., Lee, H., Hoggan, M., and Johnson, F. (1979). Adenovirus-associated Virus Structural Protein Sequence Homology. J Gen Virol *45*, 209–216.

Lukashchuk, V., and Everett, R.D. (2010). Regulation of ICP0-null mutant herpes simplex virus type 1 infection by ND10 components ATRX and hDaxx. J. Virol. *84*, 4026–4040.

Lusby, E.W., and Berns, K.I. (1982). Mapping of the 5' termini of two adeno-associated virus 2 RNAs in the left half of the genome. J. Virol. *41*, 518–526.

Lusby, E., Fife, K.H., and Berns, K.I. (1980). Nucleotide sequence of the inverted terminal repetition in adeno- associated virus DNA. J Virol *34*, 402–409.

Lusby, E., Bohenzky, R., and Berns, K.I. (1981). Inverted terminal repetition in adeno-associated virus DNA: independence of the orientation at either end of the genome. J. Virol. *37*, 1083–1086.

Malik, S., Guermah, M., and Roeder, R.G. (1998). A dynamic model for PC4 coactivator function in RNA polymerase II transcription. Proc. Natl. Acad. Sci. U. S. A. *95*, 2192–2197.

Mancebo, H.S.Y., Lee, G., Flygare, J., Tomassini, J., Luu, P., Zhu, Y., Peng, J., Blau, C., Hazuda, D., Price, D., et al. (1997). P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. Genes Dev. *11*, 2633–2644.

Manno, C.S., Pierce, G.F., Arruda, V.R., Glader, B., Ragni, M., Rasko, J.J., Rasko, J.J., Ozelo, M.C., Hoots, K., Blatt, P., et al. (2006). Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. Nat. Med. *12*, 342–347.

van Mansfeld, A.D.M., van Teeffelen, H.A.A.M., Bass, P.D., and Jansz, H.S. (1986). Two juxtaposed tyrosyl-OH groups participate in 0X174 gene a protein catalysed cleavage and ligation of DNA. Nucleic Acids Res. *14*, 4229–4238.

Mansilla-Soto, J., Yoon-Robarts, M., Rice, W.J., Arya, S., Escalante, C.R., and Linden, R.M. (2009). DNA structure modulates the oligomerization properties of the AAV initiator protein Rep68. PLoS Pathog. *5*.

Mao, Y.S., Zhang, B., and Spector, D.L. (2011). Biogenesis and function of nuclear bodies. Trends Genet. *27*, 295–306.

Marcus, C.J., Laughlin, C.A., and Carter, B.J. (1981). Adeno-associated virus RNA transcription in vivo. Eur J Biochem *121*, 147–154.

Marcus-Sekura, C.J., and Carter, B.J. (1983). Chromatin-like structure of adeno-associated virus DNA in infected cells. J Virol *48*, 79–87.

Margolin, J.F., Friedman, J.R., Meyer, W.K., Vissing, H., Thiesen, H.J., and Rauscher, F.J. (1994). Krüppel-associated boxes are potent transcriptional repression domains. Proc. Natl. Acad. Sci. U. S. A. *91*, 4509–4513.

Maroto, B., Ramírez, J.C., and Almendral, J.M. (2000). Phosphorylation status of the parvovirus minute virus of mice particle: mapping and biological relevance of the major phosphorylation sites. J. Virol. *74*, 10892–10902.

Maroto, B., Valle, N., Saffrich, R., and Almendral, J. (2004). Nuclear export of the non-enveloped parvovirus virion is directed by an unordered protein signal exposed on the capsid surface. J Virol *78*, 10694–12004.

Mascle, X.H., Germain-Desprez, D., Huynh, P., Estephan, P., and Aubry, M. (2007). Sumoylation of the transcriptional intermediary factor 1 gamma (TIF1gamma), the Co-repressor of the KRAB multifinger proteins, is required for its transcriptional activity and is modulated by the KRAB domain. J. Biol. Chem. *282*, 10190–10202.

Matsumoto, K., Nagata, K., Ui, M., and Hanaoka, F. (1993). Template activating factor I, a novel host factor required to stimulate the adenovirus core DNA replication. J. Biol. Chem. *268*, 10582–10587.

Mayor, H., and Melnick, J. (1966). Small deoxyribonucleic acid-containing viruses (picodnavirus group). Nature *210*, 331–332.

Mayor, H.D., Jamison, R.M., Jordan, L.E., and Melnick, J.L. (1965). Structure and Composition of a Small Particle Prepared from a Simian Adenovirus. J. Bacteriol. *90*, 235–242.

Mayor, H.D., Drake, S., Stahmann, J., and Mumford, D.M. (1976). Antibodies to adeno-associated satellite virus and herpes simplex in sera from cancer patients and normal adults. Am. J. Obstet. Gynecol. *126*, 100–104.

McCarty, D.M., Pereira, D.J., Zolotukhin, I., Zhou, X., Ryan, J.H., and Muzyczka, N. (1994). Identification of linear DNA sequences that specifically bind the adeno-associated virus Rep protein. J Virol *68*, 4988–4997.

McLaughlin, S.K., Collis, P., Hermonat, P.L., and Muzyczka, N. (1988). Adeno-associated virus general transduction vectors: analysis of proviral structures. J. Virol. *62*, 1963–1973.

McNamara, R.P., Reeder, J.E., McMillan, E.A., Bacon, C.W., McCann, J.L., and D'Orso, I. (2016). KAP1 Recruitment of the 7SK snRNP Complex to Promoters Enables Transcription Elongation by RNA Polymerase II. Mol. Cell *61*, 39–53.

Melendez, A., Li, W., and Kalderon, D. (1995). Activity, expression and function of a second Drosophila protein kinase A catalytic subunit gene. Genetics *141*, 1507–1520.

Mendelson, E., Trempe, J.P., and Carter, B.J. (1986). Identification of the trans-acting Rep

proteins of adeno-associated virus by antibodies to a synthetic oligopeptide. J. Virol. *60*, 823–832.

Miao, C.H., Nakai, H., Thompson, A.R., Storm, T.A., Chiu, W., Snyder, R.O., and Kay, M.A. (2000). Nonrandom transduction of recombinant adeno-associated virus vectors in mouse hepatocytes in vivo: cell cycling does not influence hepatocyte transduction. J. Virol. *74*, 3793–3803.

Miller, E.B., Gurda-Whitaker, B., Govindasamy, L., McKenna, R., Zolotukhin, S., Muzyczka, N., and Agbandje-McKenna, M. (2006). Production, purification and preliminary X-ray crystallographic studies of adeno-associated virus serotype 1. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. *62*, 1271–1274.

Mingozzi, F., Maus, M. V, Hui, D.J., Sabatino, D.E., Murphy, S.L., Rasko, J.E.J., Ragni, M. V, Manno, C.S., Sommer, J., Jiang, H., et al. (2007). CD8(+) T-cell responses to adeno-associated virus capsid in humans. Nat. Med. *13*, 419–422.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Sci. (New York, NY) *303*, 343–348.

Moggs, J.G., Grandi, P., Quivy, J.P., Jónsson, Z.O., Hübscher, U., Becker, P.B., and Almouzni, G. (2000). A CAF-1-PCNA-mediated chromatin assembly pathway triggered by sensing DNA damage. Mol. Cell. Biol. *20*, 1206–1218.

Moosmann, P., Georgiev, O., Le Douarin, B., Bourquin, J.P., and Schaffner, W. (1996). Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. Nucleic Acids Res *24*, 4859–4867.

Morgan, H.D., Santos, F., Green, K., Dean, W., and Reik, W. (2005). Epigenetic reprogramming in mammals. Hum. Mol. Genet. *14*.

Morita, E., Arii, J., Christensen, D., Votteler, J., and Sundquist, W.I. (2012). Attenuated protein expression vectors for use in siRNA rescue experiments. Biotechniques 1–5.

Mouw, M.B., and Pintel, D.J. (2000). Adeno-associated virus RNAs appear in a temporal order and their splicing is stimulated during coinfection with adenovirus. J. Virol. *74*, 9878–9888.

Myers, M.W., and Carter, B.J. (1980). Assembly of adeno-associated virus. Virology *102*, 71–82.

Nada, S., and Trempe, J.P. (2002). Characterization of adeno-associated virus rep protein inhibition of adenovirus E2a gene expression. Virology *293*, 345–355.

Nam, H.-J., Lane, M.D., Padron, E., Gurda, B., McKenna, R., Kohlbrenner, E., Aslanidi, G., Byrne, B., Muzyczka, N., Zolotukhin, S., et al. (2007). Structure of adeno-associated virus serotype 8, a gene therapy vector. J. Virol. *81*, 12260–12271.

Nam, H.-J., Gurda, B.L., McKenna, R., Potter, M., Byrne, B., Salganik, M., Muzyczka, N., and Agbandje-McKenna, M. (2011). Structural Studies of Adeno-Associated Virus Serotype 8 Capsid Transitions Associated with Endosomal Trafficking. J. Virol. *85*, 11791–1179944.

Narayanan, A., Ruyechan, W.T., and Kristie, T.M. (2007). The coactivator host cell factor-1 mediates Set1 and MLL1 H3K4 trimethylation at herpesvirus immediate early promoters for initiation of infection. Proc Natl Acad Sci U S A *104*, 10835–10840.

Nash, K., Chen, W., McDonald, W.F., Zhou, X., and Muzyczka, N. (2007). Purification of host cell enzymes involved in adeno-associated virus DNA replication. J. Virol. *81*, 5777–5787.

Nash, K., Chen, W., and Muzyczka, N. (2008). Complete in vitro reconstitution of adeno-associated virus DNA replication requires the minichromosome maintenance complex proteins. J. Virol. *82*, 1458–1464.

Nash, K., Chen, W., Salganik, M., and Muzyczka, N. (2009). Identification of cellular proteins that interact with the adeno-associated virus rep protein. J Virol *83*, 454–469.

Nayak, R., and Pintel, D.J. (2007). Adeno-associated viruses can induce phosphorylation of eIF2alpha via PKR activation, which can be overcome by helper adenovirus type 5 virus-associated RNA. J. Virol. *81*, 11908–11916.

Needham, P.G., Casper, J.M., Kalman-Maltese, V., Verrill, K., Dignam, J.D., and Trempe, J.P. (2006). Adeno-Associated Virus Rep Protein-Mediated Inhibition of Transcription of the Adenovirus Major Late Promoter In Vitro. J. Virol. *80*, 6207–6217.

Ng, R., Govindasamy, L., Gurda, B.L., McKenna, R., Kozyreva, O.G., Samulski, R.J., Parent, K.N., Baker, T.S., and Agbandje-McKenna, M. (2010). Structural characterization of the dual glycan binding adeno-associated virus serotype 6. J. Virol. *84*, 12945–12957.

Nicolas, A., Alazard-Dany, N., Biollay, C., Arata, L., Jolinon, N., Kuhn, L., Ferro, M., Weller, S.K., Epstein, A.L., Salvetti, A., et al. (2010). Identification of rep-associated factors in herpes simplex virus type 1-induced adeno-associated virus type 2 replication compartments. J. Virol. *84*, 8871–8887.

Nicolas, A., Jolinon, N., Alazard-Dany, N., Barateau, V., Epstein, A.L., Greco, A., Büning, H., and Salvetti, A. (2012). Factors influencing helper-independent adeno-associated virus replication. Virology *432*, 1–9.

Nielsen, A.L., Ortiz, J.A., You, J., Oulad-Abdelghani, M., Khechumian, R., Gansmuller, A., Chambon, P., and Losson, R. (1999). Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. EMBO J *18*, 6385–6395.

Niwa, O., Yokota, Y., Ishida, H., and Sugahara, T. (1983). Independent mechanisms involved in suppression of the moloney leukemia virus genome during differentiation of murine teratocarcinoma cells. Cell *32*, 1105–1113.

Nokta, M.A., and Pollard, R.B. (1992). Human immunodeficiency virus replication: modulation by cellular levels of cAMP. AIDS Res. Hum. Retroviruses *8*, 1255–1261.

Nonnenmacher, M., and Weber, T. (2011). Adeno-associated virus 2 infection requires endocytosis through the CLIC/GEEC pathway. Cell Host Microbe *10*, 563–576.

Noon, A.T., Shibata, A., Rief, N., Lobrich, M., Stewart, G.S., Jeggo, P.A., and Goodarzi, A.A. (2010). 53BP1-dependent robust localized KAP-1 phosphorylation is essential for heterochromatic DNA double-strand break repair. Nat Cell Biol *12*, 177–184.

O'Connell, N., Nichols, S.R., Heroes, E., Beullens, M., Bollen, M., Peti, W., and Page, R. (2012). The molecular basis for substrate specificity of the nuclear NIPP1:PP1 holoenzyme. Structure *20*, 1746–1756.

O'Geen, H., Squazzo, S.L., Iyengar, S., Blahnik, K., Rinn, J.L., Chang, H.Y., Green, R., and Farnham, P.J. (2007). Genome-wide analysis of KAP1 binding suggests autoregulation of KRAB-ZNFs. PLoS Genet *3*, e89.

Odegrip, R., and Haggård-Ljungquist, E. (2001). The two active-site tyrosine residues of the a protein play non-equivalent roles during initiation of rolling circle replication of bacteriophage p2. J. Mol. Biol. *308*, 147–163.

Odunsi, K.O., van Ee, C.C., Ganesan, T.S., and Shelling, a N. (2000). Evaluation of the possible protective role of adeno-associated virus type 2 infection in HPV-associated premalignant disease of the cervix. Gynecol. Oncol. *78*, 342–345.

Oh, J., Ruskoski, N., and Fraser, N.W. (2012). Chromatin assembly on herpes simplex virus 1 DNA early during a lytic infection is Asf1a dependent. J. Virol. *86*, 12313–12321.

Okoro, D.R., Arva, N., Gao, C., Polotskaia, A., Puente, C., Rosso, M., and Bargonetti, J. (2013). Endogenous Human MDM2-C Is Highly Expressed in Human Cancers and Functions as a p53-Independent Growth Activator. PLoS One *8*.

Okuwaki, M., Iwamatsu, a, Tsujimoto, M., and Nagata, K. (2001). Identification of nucleophosmin/B23, an acidic nucleolar protein, as a stimulatory factor for in vitro replication of adenovirus DNA complexed with viral basic core proteins. J. Mol. Biol. *311*, 41–55.

Opie, S.R., Warrington Jr., K.H., Agbandje-McKenna, M., Zolotukhin, S., and Muzyczka, N. (2003). Identification of amino acid residues in the capsid proteins of adeno-associated virus type 2 that contribute to heparan sulfate proteoglycan binding. J Virol *77*, 6995–7006.

Orazio, N.I., Naeger, C.M., Karlseder, J., and Weitzman, M.D. (2011). The adenovirus E1b55K/E4orf6 complex induces degradation of the Bloom helicase during infection. J. Virol. *85*, 1887–1892.

Orzalli, M., DeLuca, N., and Knipe, D. (2012). Nuclear IFI16 induction of IRF-3 signaling during herpesviral infection and degradation of IFI16 by the viral ICP0 protein. PNAS *109*, 3008–3017.

Owens, R. a, Weitzman, M.D., Kyöstiö, S.R., and Carter, B.J. (1993). Identification of a DNA-binding domain in the amino terminus of adeno-associated virus Rep proteins. J. Virol. *67*, 997–1005.

Ozato, K., Shin, D.-M., Chang, T.-H., and Morse, H.C. (2008). TRIM family proteins and their emerging roles in innate immunity. Nat. Rev. Immunol. *8*, 849–860.

Padron, E., Bowman, V., Kaludov, N., Govindasamy, L., Levy, H., Nick, P., McKenna, R., Muzyczka, N., Chiorini, J.A., Baker, T.S., et al. (2005). Structure of adeno-associated virus type 4. J. Virol. *79*, 5047–5058.

Pagano, J.M., Kwak, H., Waters, C.T., Sprouse, R.O., White, B.S., Ozer, A., Szeto, K., Shalloway, D., Craighead, H.G., and Lis, J.T. (2014). Defining NELF-E RNA Binding in HIV-1 and Promoter-Proximal Pause Regions. PLoS Genet. *10*.

Paganon, M., Durst, M., Joswig, SDraetta, G., and Jansen-Durr, P. (1992). Binding of the human E2F transription factor to the retinoblastoma protein but not to cyclin A is abolished in HPV-16 immortalized cells. Oncogene *7*, 1681–1686.

Pasquale, G.D.I., Stacey, S.N., and Pasquale, G.D.I. (1998). Adeno-Associated Virus Rep78 Protein Interacts with Protein Kinase A and Its Homolog PRKX and Inhibits CREB-Dependent Transcriptional Adeno-Associated Virus Rep78 Protein Interacts with Protein Kinase A and Its Homolog PRKX and Inhibits CREB- Dependent Tr. *72*, 7916–7925.

Di Pasquale, G., and Chiorini, J.A. (2006). AAV transcytosis through barrier epithelia and endothelium. Mol. Ther. *13*, 506–516.

Di Pasquale, G., and Stacey, S.N. (1998). Adeno-associated virus Rep78 protein interacts with protein kinase A and its homolog PRKX and inhibits CREB-dependent transcriptional activation. J Virol *72*, 7916–7925.

Di Pasquale, G., Davidson, B.L., Stein, C.S., Martins, I., Scudiero, D., Monks, A., and Chiorini, J. a (2003). Identification of PDGFR as a receptor for AAV-5 transduction. Nat. Med. *9*, 1306–1312.

Pearson, J.L., Robinson, T.J., Muoz, M.J., Kornblihtt, A.R., and Garcia-Blanco, M.A. (2008). Identification of the cellular targets of the transcription factor TCERG1 reveals a prevalent role in mrna processing. J. Biol. Chem. *283*, 7949–7961.

Pegoraro, G., Marcello, A., Myers, M.P., and Giacca, M. (2006). Regulation of adeno-associated virus DNA replication by the cellular TAF-I/set complex. J Virol *80*, 6855–6864.

Pei, R., and Berk, A.J. (1989). Multiple transcription factor binding sites mediate adenovirus E1A transactivation. J. Virol. *63*, 3499–3506.

Penaud-Budloo, M., Le Guiner, C., Nowrouzi, A., Toromanoff, A., Cherel, Y., Chenuaud, P., Schmidt, M., von Kalle, C., Rolling, F., Moullier, P., et al. (2008). Adeno-associated virus vector genomes persist as episomal chromatin in primate muscle. J Virol *82*, 7875–7885.

Peng, H., Begg, G.E., Schultz, D.C., Friedman, J.R., Jensen, D.E., Speicher, D.W., and Rauscher, F.J. (2000). Reconstitution of the KRAB-KAP-1 repressor complex: a model system for defining the molecular anatomy of RING-B box-coiled-coil domain-mediated protein-protein interactions. J. Mol. Biol. *295*, 1139–1162.

Peng, H., Feldman, I., and Rauscher, F.J. (2002). Hetero-oligomerization among the TIF family of RBCC/TRIM domain-containing nuclear cofactors: A potential mechanism for regulating the switch between coactivation and corepression. J. Mol. Biol. *320*, 629–644.

Peng, H., Ivanov, A. V., Oh, H.J., Lau, Y.F.C., and Rauscher, F.J. (2009). Epigenetic gene silencing by the SRY protein is mediated by a KRAB-O protein that recruits the KAP1 co-repressor machinery. J. Biol. Chem. *284*, 35670–35680.

Peng, J., Liu, M., Marion, J., Zhu, Y., and Price, D.H. (1998). RNA polymerase II elongation control. In Cold Spring Harbor Symposia on Quantitative Biology, pp. 365–370.

Pereira, D.J., and Muzyczka, N. (1997a). The adeno-associated virus type 2 p40 promoter requires a proximal Sp1 interaction and a p19 CArG-like element to facilitate Rep transactivation. J. Virol. *71*, 4300–4309.

Pereira, D.J., and Muzyczka, N. (1997b). The cellular transcription factor SP1 and an unknown cellular protein are required to mediate Rep protein activation of the adeno-associated virus p19 promoter. J Virol *71*, 1747–1756.

Pereira, D.J., McCarty, D.M., and Muzyczka, N. (1997). The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection. J Virol *71*, 1079–1088.

Perros, M., Deleu, L., Vanacker, J., Kherrouche, Z., Spruyt, N., Faisst, S., and Rommelaere, J. (1995). Upstream CREs participate in the basal activity of minute virus of mice promoter P4 and its stimulation in ras transformed cells. J Virol *69*, 5506–5515.

Peterlin, B.M., and Price, D.H. (2006). Controlling the Elongation Phase of Transcription with P-TEFb. Mol. Cell *23*, 297–305.

Petri, K., Gabriel, R., Agundez, L., Fronza, R., Afzal, S., Kaeppel, C., Linden, R.M., Henckaerts, E., and Schmidt, M. (2015). Presence of a trs-Like Motif Promotes Rep-Mediated Wild-Type Adeno-Associated Virus Type 2 Integration. J. Virol. *89*, 7428–7432.

Philpott, N.J., Giraud-Wali, C., Dupuis, C., Gomos, J., Hamilton, H., Berns, K.I., and Falck-Pedersen, E. (2002). Efficient integration of recombinant adeno-associated virus DNA vectors requires a p5-rep sequence in cis. J. Virol. *76*, 5411–5421.

Pignattii, P.F., and Cassai, E. (1980). Analysis of Herpes Simplex Virus Nucleoprotein Complexes Extracted from Infected Cells. J. Virol. *36*, 816–828.

Pil, P.M., and Lippard, S.J. (1992). Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin. Science *256*, 234–237.

Pillay, S., Meyer, N.L., Puschnik, A.S., Davulcu, O., Diep, J., Ishikawa, Y., Jae, L.T., Wosen, J.E., Nagamine, C.M., Chapman, M.S., et al. (2016). An essential receptor for adeno-associated virus infection. Nature *530*, 108–112.

Poole, E., and Sinclair, J. (2015). Sleepless latency of human cytomegalovirus. Med. Microbiol. Immunol. *204*, 421–429.

Popa-Wagner, R., Sonntag, F., Schmidt, K., King, J., and Kleinschmidt, J. a (2012). Nuclear translocation of adeno-associated virus type 2 capsid proteins for virion assembly. J. Gen. Virol. *93*, 1887–1898.

Prasad, K.M., and Trempe, J.P. (1995). The adeno-associated virus Rep78 protein is covalently linked to viral DNA in a preformed virion. Virology *214*, 360–370.

Prasad, C.K., Meyers, C., Zhan, D.J., You, H., Chiriva-Internati, M., Mehta, J.L., Liu, Y., and Hermonat, P.L. (2003). The adeno-associated virus major regulatory protein Rep78-c-Jun-DNA motif complex modulates AP-1 activity. Virology *314*, 423–431.

Prasad, K.M., Zhou, C., and Trempe, J.P. (1997). Characterization of the Rep78/adeno-associated virus complex. Virology *229*, 183–192.

Qiao, C., Yuan, Z., Li, J., He, B., Zheng, H., Mayer, C., and Xiao, X. (2011). Liver-specific microRNA-122 target sequences incorporated in AAV vectors efficiently inhibits transgene expression in the liver. Gene Ther. *18*, 403–410.

Qing, K., Mah, C., Hansen, J., Zhou, S., Dwarki, V., and Srivastava, A. (1999). Human fibroblast growth factor receptor 1 is a co-receptor for infection by adeno-associated virus 2. Nat Med *5*, 71–77.

Qiu, J., and Pintel, D.J. (2002). The adeno-associated virus type 2 Rep protein regulates RNA processing via interaction with the transcription template. Mol. Cell. Biol. *22*, 3639–3652.

Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., Baglivo, I., Pedone, P. V., Grimaldi, G., Riccio, A., et al. (2011). In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. Mol. Cell *44*, 361–372.

Querido, E., Blanchette, P., Yan, Q., Kamura, T., Morrison, M., Boivin, D., Kaelin, W.G., Conaway, R.C., Conaway, J.W., and Branton, P.E. (2001). Degradation of p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex. Genes Dev. *15*, 3104–3117.

Quesada, O., Gurda, B., Govindasamy, L., McKenna, R., Kohlbrenner, E., Aslanidi, G., Zolotukhin, S., Muzyczka, N., and Agbandje-Mckenna, M. (2007). Production, purification and preliminary X-ray crystallographic studies of adeno-associated virus serotype 7. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. *63*, 1073–1076.

Rasaiyaah, J., Tan, C.P., Fletcher, A.J., Price, A.J., Blondeau, C., Hilditch, L., Jacques, D. a, Selwood, D.L., James, L.C., Noursadeghi, M., et al. (2013). HIV-1 evades innate immune recognition through specific cofactor recruitment. Nature *503*, 402–405.

Rauwel, B., Jang, S.M., Cassano, M., Kapopoulou, A., Barde, I., and Trono, D. (2015).

Release of human cytomegalovirus from latency by a KAP1/TRIM28 phosphorylation switch. Elife *4*.

Redemann, B.E., Mendelson, E., and Carter, B.J. (1989). Adeno-associated virus rep protein synthesis during productive infection. J. Virol. *63*, 873–882.

Rose, J., and Koczot, F. (1971). Adenovirus-associated virus multiplication. J Virol *8*, 771–777

Rose, J., Hoggan, M., and Shatkin, A. (1966). Nucleiic acid from an adeno-assoicated virus: chemical and phsyical properties. Proc Natl Acad Sci U S A *56*, 86–92.

Rose, J., Berns, K., Hoggan, M., and Koczot, F. (1969). Evidence for a single-stranded adenovirus-associated virus genome: formation of a DNA denisty hybrid on release of viral DNA. Proc Natl Acad Sci U S A *64*, 863–869.

Rose, J. a, Maizel, J. V, Inman, J.K., and Shatkin, a J. (1971). Structural proteins of adenovirus-associated viruses. J. Virol. *8*, 766–770.

Rosenbaum, M., Edwards, E., Pierce, W., Peckinpaugh, R., Parks, W., and Melnick, J. (1971). Serologic surveillance for adeno-associated satellite virus antibody in military recruits. J Immunol *106*, 711–720.

Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J Cell Biol *196*, 801–810.

Rowe, H.M., and Trono, D. (2011). Dynamic control of endogenous retroviruses during development. Virology *411*, 273–287.

Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P. V, Layard-Liesching, H., Verp, S., Marquis, J., et al. (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. Nature *463*, 237–240.

Rowe, H.M., Friedli, M., Offner, S., Verp, S., Mesnard, D., Marquis, J., Aktas, T., and Trono, D. (2013). De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET. Development *140*, 519–529.

Rueegsegger, U., Blank, D., and Keller, W. (1998). Human pre-mRNA cleavage factor Im is related to spliceosomal SR proteins and can be reconstituted in vitro from recombinant subunits. Mol Cell *1*, 243--253.

Ryan, R.F., Schultz, D.C., Ayyanathan, K., Singh, P.B., Friedman, J.R., Fredericks, W.J., Rauscher 3rd, F.J., and Rauscher, F.J. (1999). KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. Mol Cell Biol *19*, 4366–4378.

Salganik, M., Aydemir, F., Nam, H.-J., McKenna, R., Agbandje-McKenna, M., and Muzyczka, N. (2014). Adeno-associated virus capsid proteins may play a role in transcription and second-strand synthesis of recombinant genomes. J. Virol. *88*, 1071–1079.

Samulski, R.J., and Shenk, T. (1988). Adenovirus E1B 55-Mr polypeptide facilitates timely cytoplasmic accumulation of adeno-associated virus mRNAs. J. Virol. *62*, 206–210.

Samulski, R.J., Berns, K.I., Tan, M., and Muzyczka, N. (1982). Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. Proc. Natl. Acad. Sci. U. S. A. *79*, 2077–2081.

Samulski, R.J., Srivastava, A., Berns, K.I., and Muzyczka, N. (1983). Rescue of adeno-

associated virus from recombinant plasmids: Gene correction within the terminal repeats of AAV. Cell *33*, 135–143.

Samulski, R.J., Chang, L.S., and Shenk, T. (1989). Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. J. Virol. *63*, 3822–3828.

Samulski, R.J., Zhu, X., Xiao, X., Brook, J.D., Housman, D.E., Epstein, N., and Hunter, L.A. (1991). Targeted integration of adeno-associated virus (AAV) into human chromosome 19. EMBO J *10*, 3941–3950.

Sanlioglu, S., Benson, P.K., Yang, J., Atkinson, E.M., Reynolds, T., and Engelhardt, J.F. (2000). Endocytosis and nuclear trafficking of adeno-associated virus type 2 are controlled by rac1 and phosphatidylinositol-3 kinase activation. J. Virol. *74*, 9184–9196.

Saudan, P., Vlach, J., and Beard, P. (2000). Inhibition of S-phase progression by adeno-associated virus Rep78 protein is mediated by hypophosphorylated pRb. EMBO J *19*, 4351–4361.

Saurin, A.J., Borden, K.L.B., Boddy, M.N., and Freemont, P.S. (1996). Does this have a familiar RING? Trends Biochem. Sci. *21*, 208–214.

Schlesinger, S., Lee, A.H., Wang, G.Z., Green, L., and Goff, S.P. (2013). Proviral silencing in embryonic cells is regulated by Yin Yang 1. Cell Rep *4*, 50–58.

Schmidt, M., Afione, S., and Kotin, R.M. (2000). Adeno-associated virus type 2 Rep78 induces apoptosis through caspase activation independently of p53. J. Virol. *74*, 9441–9450.
Schnepp, B.C., Clark, K.R., Klemanski, D.L., Pacak, C.A., and Johnson, P.R. (2003). Genetic fate of recombinant adeno-associated virus vector genomes in muscle. J. Virol. *77*, 3495–3504.

Schnepp, B.C., Jensen, R.L., Chen, C.L., Johnson, P.R., and Clark, K.R. (2005). Characterization of adeno-associated virus genomes isolated from human tissues. J Virol *79*, 14793–14803.

Schreiner, S., Wimmer, P., Sirma, H., Everett, R.D., Blanchette, P., Groitl, P., and Dobner, T. (2010). Proteasome-dependent degradation of Daxx by the viral E1B-55K protein in human adenovirus-infected cells. J. Virol. *84*, 7029–7038.

Schreiner, S., Kinkley, S., Burck, C., Mund, A., Wimmer, P., Schubert, T., Groitl, P., Will, H., and Dobner, T. (2013a). SPOC1-mediated antiviral host cell response is antagonized early in human adenovirus type 5 infection. PLoS Pathog *9*, e1003775.

Schreiner, S., Bürck, C., Glass, M., Groitl, P., Wimmer, P., Kinkley, S., Mund, A., Everett, R.D., and Dobner, T. (2013b). Control of human adenovirus type 5 gene expression by cellular Daxx/ATRX chromatin-associated complexes. Nucleic Acids Res. *41*, 3532–3550.

Schultz, D.C., Friedman, J.R., and Rauscher 3rd, F.J. (2001). Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. Genes Dev *15*, 428–443.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher 3rd, F.J. (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev *16*, 919–932.

Schwartz, R.A., Palacios, J.A., Cassell, G.D., Adam, S., Giacca, M., and Weitzman, M.D. (2007). The Mre11/Rad50/Nbs1 complex limits adeno-associated virus transduction and replication. J Virol *81*, 12936–12945.

Schwartz, R.A., Carson, C.T., Schuberth, C., and Weitzman, M.D. (2009). Adeno-associated virus replication induces a DNA damage response coordinated by DNA-dependent protein kinase. J Virol *83*, 6269–6278.

Seki, Y., Kurisaki, A., Watanabe-Susaki, K., Nakajima, Y., Nakanishi, M., Arai, Y., Shiota, K., Sugino, H., and Asashima, M. (2010). TIF1β regulates the pluripotency of embryonic stem cells in a phosphorylation-dependent manner. Proc Natl Acad Sci U S A *107*, 10926–10931.

Seto, E., Shi, Y., and Shenk, T. (1991). YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro. Nature *354*, 241–245.

Shelling, A.N., and Smith, M.G. (1994). Targeted integration of transfected and infected adeno-associated virus vectors containing the neomycin resistance gene. Gene Ther *1*, 165–169.

Shen, S., Bryant, K.D., Brown, S.M., Randell, S.H., and Asokan, A. (2011). Terminal n-linked galactose is the primary receptor for adeno-associated virus. J. Biol. Chem. *286*, 13532–13540.

Shen, X., Mizuguchi, G., Hamiche, a, and Wu, C. (2000). A chromatin remodelling complex involved in transcription and DNA processing. Nature *406*, 541–544.

Shi, Y., Seto, E., Chang, L.S., and Shenk, T. (1991). Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein. Cell *67*, 377–388.

Silva, F.P., Hamamoto, R., Furukawa, Y., and Nakamura, Y. (2006). TIPUH1 encodes a novel KRAB zinc-finger protein highly expressed in human hepatocellular carcinomas. Oncogene *25*, 5063–5070.

Singh, V.V., Kerur, N., Bottero, V., Dutta, S., Chakraborty, S., Ansari, M.A., Paudel, N., Chikoti, L., and Chandran, B. (2013). Kaposi's sarcoma-associated herpesvirus latency in endothelial and B cells activates gamma interferon-inducible protein 16-mediated inflammasomes. J. Virol. *87*, 4417–4431.

Smale, S.T. (2010). Selective Transcription in Response to an Inflammatory Stimulus. Cell *140*, 833–844.

Smiley, J.R. (2004) Herpes simplex virus virion host shutoff protein: immune evasion mediated by a viral RNase? J. Virol. *78,* 1063-1068.

Smith, R.H., and Kotin, R.M. (1998). The Rep52 gene product of adeno-associated virus is a DNA helicase with 3'-to-5' polarity. J. Virol. *72*, 4874–4881.

Smith, R.H., and Kotin, R.M. (2000). An adeno-associated virus (AAV) initiator protein, Rep78, catalyzes the cleavage and ligation of single-stranded AAV ori DNA. J. Virol. *74*, 3122–3129.

Smith, S., and Stillman, B. (1989). Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. Cell *58*, 15–25.

Smith, D., Ward, P., and Linden, R. (1999). Comparative characterization of Rep proteins from the helper-dependent adeno-associated virus type 2 and the autonomous goose parvovirus. J Virol *73*, 2930–2937.

Smith, J., Herrero, R., Erles, K., Grimm, D., Munoz, N., Bosch, F.X., Tafur, L., Shah, K. V., and Schlehofer, J.R. (2001). Adeno-associated virus seropositivity and HPV-induced cervical cancer in Spain and Colombia. Int. J. Cancer *94*, 520–526.

Smith, R.H., Spano, a J., and Kotin, R.M. (1997). The Rep78 gene product of adeno-associated virus (AAV) self-associates to form a hexameric complex in the presence of AAV ori sequences. J. Virol. *71*, 4461–4471.

Snyder, R.O., Im, D.S., Ni, T., Xiao, X., Samulski, R.J., and Muzyczka, N. (1993). Features of the adeno-associated virus origin involved in substrate recognition by the viral Rep protein. J. Virol. *67*, 6096–6104.

Sonntag, F., Bleker, S., Leuchs, B., Fischer, R., and Kleinschmidt, J.A. (2006). Adeno-associated virus type 2 capsids with externalized VP1/VP2 trafficking domains are generated prior to passage through the cytoplasm and are maintained until uncoating occurs in the nucleus. J. Virol. *80*, 11040–11054.

Sonntag, F., Schmidt, K., and Kleinschmidt, J.A. (2010). A viral assembly factor promotes AAV2 capsid formation in the nucleolus. Proc. Natl. Acad. Sci. U. S. A. *107*, 10220–10225.

Spector, D.J. (2007). Default assembly of early adenovirus chromatin. Virology *359*, 116–125.

Sripathy, S.P., Stevens, J., and Schultz, D.C. (2006). The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. Mol. Cell. Biol. *26*, 8623–8638.

Srivastava, A., Lusby, E.W., and Berns, K.I. (1983). Nucleotide sequence and organization of the adeno-associated virus 2 genome. J. Virol. *45*, 555–564.

Stahnke, S., Lux, K., Uhrig, S., Kreppel, F., Hösel, M., Coutelle, O., Ogris, M., Hallek, M., and Buning, H. (2011). Intrinsic phospholipase A2 activity of adeno-associated virus is involved in endosomal escape of incoming particles. Virology *409*, 77–83.

Stillman, B. (1986). Chromatin assembly during SV40 DNA replication in vitro. Cell *45*, 555–565.

Stracker, T.H., Carson, C.T., and Weitzman, M.D. (2002). Adenovirus oncoproteins inactivate the Mre11 Rad50 NBS1 DNA repair complex. Nature *418*, 348.

Stracker, T.H., Cassell, G.D., Ward, P., Loo, Y., van Breukelen, B., Carrington-Lawrence, S.D., Hamatake, R.K., van der Vliet, P.C., Weller, S.K., Melendy, T., et al. (2004). The Rep protein of adeno-associated virus type 2 interacts with single-stranded DNA-binding proteins that enhance viral replication. J. Virol. *78*, 441–453.

Stratmann, S., Morrone, S., van Oijen, A., and Sohn, J. (2015). The innate immune sensor IFI16 recognizes foreign DNA in the nucleus by scanning along the duplex. Elife *4:e11721*.

Straus, S., Sebring, E., and Rose, J. (1978). Replication of Mammalian Parvoviruses.

Straus, S.E., Sebring, E.D., and Rose, J.A. (1976). Concatemers of alternating plus and minus strands are intermediates in adenovirus-associated virus DNA synthesis. Biochem. Proc. Natl. Acad. Sci. United States Am. *73*, 742–746.

Strickler, H.D., Viscidi, R., Escoffery, C., Rattray, C., Kotloff, K.L., Goldberg, J., Manns, A., Rabkin, C., Daniel, R., Hanchard, B., et al. (1999). Adeno-associated virus and development of cervical neoplasia. J. Med. Virol. *59*, 60–65.

Summerford, C., and Samulski, R.J. (1998). Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. J Virol *72*, 1438–1445.

Summerford, C., Bartlett, J.S., and Samulski, R.J. (1999). AlphaVbeta5 integrin: a co-receptor for adeno-associated virus type 2 infection. Nat. Med. *5*, 78–82.

Sun, R., Liang, D., Gao, Y., and Lan, K. (2014). Kaposi's sarcoma-associated herpesvirus-encoded LANA interacts with host KAP1 to facilitate establishment of viral latency. J Virol *88*, 7331–7344.

Sune, C., and Garcia-Blanco, M. (1999). Transcriptional Cofactor CA150 Regulates RNA Polymerase II Elongation in a TATA-Box-Dependent Manner. Mol. Cell. Biol. *19*, 4719–4728.

Sune, C., Hayashi, T., Liu, Y., Lane, W.S., Young, R. a, and Garcia-Blanco, M. (1997). CA150, a nuclear protein associated with the RNA polymerase II holoenzyme, is involved in Tat-activated human immunodeficiency virus type 1 transcription. Mol. Cell. Biol. *17*, 6029–6039.

Suomalainen, M., Nakano, M.Y., Boucke, K., Keller, S., and Greber, U.F. (2001). Adenovirus-activated PKA and p38/MAPK pathways boost microtubule-mediated nuclear targeting of virus. EMBO J. *20*, 1310–1319.

Surovsky, R., Urabe, M., Godwin, S.G., Surosky, R.T., Urabe, M., Godwin, S.G., McQuiston, S.A., Kurtzman, G.J., Ozawa, K., and Natsoulis, G. (1997). Adeno-associated virus Rep proteins target DNA sequences to a unique locus in the human genome. J Virol *71*, 7951–7959.

Tan, I., Ng, C.H., Lim, L., and Leung, T. (2001). Phosphorylation of a novel myosin binding subunit of protein phosphatase 1 reveals a conserved mechanism in the regulation of actin cytoskeleton. J. Biol. Chem. *276*, 21209–21216.

Tang, X., Hui, Z.-G., Cui, X.-L., Garg, R., Kastan, M.B., and Xu, B. (2008). A novel ATM-dependent pathway regulates protein phosphatase 1 in response to DNA damage. Mol. Cell. Biol. *28*, 2559–2566.

Tao, R.H., Kawate, H., Wu, Y., Ohnaka, K., Ishizuka, M., Inoue, A., Hagiwara, H., and Takayanagi, R. (2006). Testicular zinc finger protein recruits histone deacetylase 2 and suppresses the transactivation function and intranuclear foci formation of agonist-bound androgen receptor competitively with TIF2. Mol. Cell. Endocrinol. *247*, 150–165.

Tattersall, P. (2006). The evolution of parvovirus taxonomy. In Parvoviruses, pp. 5–14.

Tattersall, P., and Ward, D.C. (1976). Rolling hairpin model for replication of parvovirus and linear chromosomal DNA. Nature *263*, 106–109.

Tezak, Z., Nagaraju, K., Plotz, P., and Hoffman, E. (2000). Detection of adeno-associated virus in normal and myositis human skeletal muscle. Neurology *55*, 1913–1917.

Timpe, J.M., Verrill, K.C., and Trempe, J.P. (2006). Effects of adeno-associated virus on adenovirus replication and gene expression during coinfection. J. Virol. *80*, 7807–7815.

Tomimatsu, N., Mukherjee, B., and Burma, S. (2009). Distinct roles of ATR and DNA-PKcs in triggering DNA damage responses in ATM-deficient cells. EMBO Rep *10*, 629–635.

Toolan, H. (1961). A virus associated with transplantable human tumors. Bull. New York Acad. Med. *37*, 305–310.

Torreira, E., Jha, S., Lopez-Blanco, J.R., Arias-Palomo, E., Chacón, P., Cañas, C., Ayora, S., Dutta, A., and Llorca, O. (2008). Architecture of the Pontin/Reptin Complex, Essential in the Assembly of Several Macromolecular Complexes. Structure *16*, 1511–1520.

Tratschin, J.D., Miller, I.L., and Carter, B.J. (1984). Genetic analysis of adeno-associated virus: properties of deletion mutants constructed in vitro and evidence for an adeno-associated virus replication function. J. Virol. *51*, 611–619.

Tratschin, J.D., Tal, J., and Carter, B.J. (1986). Negative and positive regulation in trans of gene expression from adeno-associated virus vectors in mammalian cells by a viral rep gene product. Mol Cell Biol *6*, 2884–2894.

Trempe, J.P., and Carter, B.J. (1988a). Alternate mRNA splicing is required for synthesis of adeno-associated virus VP1 capsid protein. J. Virol. *62*, 3356–3363.

Trempe, J.P., and Carter, B.J. (1988b). Regulation of adeno-associated virus gene expression in 293 cells: control of mRNA abundance and translation. J. Virol. *62*, 68–74.

Trinkle-Mulcahy, L., Sleeman, J.E., and Lamond, a I. (2001). Dynamic targeting of protein phosphatase 1 within the nuclei of living mammalian cells. J. Cell Sci. *114*, 4219–4228.

Tsukiyama, T., Niwa, O., and Yokoro, K. (1989). Mechanism of suppression of the long terminal repeat of Moloney leukemia virus in mouse embryonal carcinoma cells. Mol. Cell. Biol. *9*, 4670–4676.

Tsuruma, R., Ohbayashi, N., Kamitani, S., Ikeda, O., Sato, N., Muromoto, R., Sekine, Y., Oritani, K., and Matsuda, T. (2008). Physical and functional interactions between STAT3 and KAP1. Oncogene *27*, 3054–3059.

Unterholzner, L., Keating, S.E., Baran, M., Horan, K.A., Jensen, S.B., Sharma, S., Sirois, C.M., Jin, T., Latz, E., Xiao, T.S., et al. (2010). IFI16 is an innate immune sensor for intracellular DNA. Nat. Immunol. *11*, 997–1004.

Urecelay, E., Ward, P., Wiener, S., Safer, B., and Kotin, R. (1995). Asymmetric replication in vitro from a human sequence element is dependent on adeno-asociated virus Rep rptoein. J Virol *69*, 2038–2046.

Urrutia, R. (2003). Protein family review KRAB-containing zinc-finger repressor proteins.

Venturini, L., You, J., Stadler, M., Galien, R., Lallemand, V., Koken, M.H., Mattei, M.G., Ganser, a, Chambon, P., Losson, R., et al. (1999). TIF1gamma, a novel member of the transcriptional intermediary factor 1 family. Oncogene *18*, 1209–1217.

Vink, E.I., Yondola, M.A., Wu, K., and Hearing, P. (2012). Adenovirus E4-ORF3-dependent relocalization of TIF1alpha and TIF1gamma relies on access to the Coiled-Coil motif. Virology *422*, 317–325.

Vogel, M.J., Guelen, L., De Wit, E., Peric-Hupkes, D., Loden, M., Talhout, W., Feenstra, M., Abbas, B., Classen, A.K., and Van Steensel, B. (2006). Human heterochromatin proteins form large domains containing KRAB-ZNF genes. Genome Res. *16*, 1493–1504.

Vogel, R., Seyffert, M., Strasser, R., de Oliveira, A.P., Dresch, C., Glauser, D.L., Jolinon, N., Salvetti, A., Weitzman, M.D., Ackermann, M., et al. (2012). Adeno-associated virus type 2 modulates the host DNA damage response induced by herpes simplex virus 1 during coinfection. J Virol *86*, 143–155.

Vogel, R., Seyffert, M., Pereira Bde, A., and Fraefel, C. (2013). Viral and Cellular Components of AAV2 Replication Compartments. Open Virol J *7*, 98–120.

Walters, R.W., Yi, S.M.P., Keshavjee, S., Brown, K.E., Welsh, M.J., Chiorini, J.A., and Zabner, J. (2001). Binding of Adeno-associated Virus Type 5 to 2,3-Linked Sialic Acid is Required for Gene Transfer. J. Biol. Chem. *276*, 20610–20616.

Walters, R.W., Agbandje-McKenna, M., Bowman, V.D., Moninger, T.O., Olson, N.H., Seiler, M., Chiorini, J.A., Baker, T.S., and Zabner, J. (2004). Structure of adeno-associated virus serotype 5. J. Virol. *78*, 3361–3371.

Wang, C., Ivanov, A., Chen, L., Fredericks, W.J., Seto, E., Rauscher 3rd, F.J., and Chen, J. (2005). MDM2 interaction with nuclear corepressor KAP1 contributes to p53 inactivation. EMBO J *24*, 3279–3290.

Wang, C., Rauscher, F.J., Cress, W.D., and Chen, J. (2007). Regulation of E2F1 function by the nuclear corepressor KAP1. J. Biol. Chem. *282*, 29902–29909.

Wang, X., Qing, K., Pnnazhagan, S., and Srivastava, A. (1997). Adeno-associated virus type 2 DNA replication in vivo mutation analyses of the D sequence in viral inverted terminal repeats. J Virol *71*, 3077–3082.

Ward, P. (2006). Replication of adeno-associated virus DNA. In Parvoviruses, pp. 189–211.

Ward, P., Dean, F.B., O'Donnell, M.E., and Berns, K.I. (1998). Role of the adenovirus DNA-binding protein in in vitro adeno-associated virus DNA replication. J Virol *72*, 420–427.

Ward, P., Falkenberg, M., Elias, P., Weitzman, M., and Linden, R.M. (2001). Rep-dependent initiation of adeno-associated virus type 2 DNA replication by a herpes simplex virus type 1 replication complex in a reconstituted system. J. Virol. *75*, 10250–10258.

Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., et al. (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520–562.

Weger, S., Wistuba, A., Grimm, D., and Kleinschmidt, J.A. (1997). Control of adeno-associated virus type 2 cap gene expression: relative influence of helper virus, terminal repeats, and Rep proteins. J. Virol. *71*, 8437–8447.

Weger, S., Wendland, M., Kleinschmidt, J.A., and Heilbronn, R. (1999). The adeno-associated virus type 2 regulatory proteins rep78 and rep68 interact with the transcriptional coactivator PC4. J Virol *73*, 260–269.

Weindler, F.W., and Heilbronn, R. (1991). A subset of herpes simplex virus replication genes provides helper functions for productive adeno-associated virus replication. J. Virol. *65*, 2476–2483.

Weitzman, M.D., Kyöstiö, S.R., Kotin, R.M., and Owens, R.A. (1994). Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. Proc. Natl. Acad. Sci. U. S. A. *91*, 5808–5812.

Weitzman, M.D., Fisher, K.J., and Wilson, J.M. (1996). Recruitment of wild-type and recombinant adeno-associated virus into adenovirus replication centers. J. Virol. *70*, 1845–1854.

Weller, M.L., Amornphimoltham, P., Schmidt, M., Wilson, P.A., Gutkind, J.S., and Chiorini, J.A. (2010). Epidermal growth factor receptor is a co-receptor for adeno-associated virus serotype 6. Nat Med *16*, 662–664.

Wen, W., Taylor, S.S., and Meinkoth, J.L. (1995). The expression and intracellular distribution of the heat-stable protein kinase inhibitor is cell cycle regulated. J. Biol. Chem. *270*, 2041–2046.

White, D.E., Negorev, D., Peng, H., Ivanov, A. V., Maul, G.G., Rauscher, F.J., and Rauscher 3rd, F.J. (2006). KAP1, a novel substrate for PIKK family members, colocalizes with numerous damage response factors at DNA lesions. Cancer Res *66*, 11594–11599.

Winocour, E., Callaham, M.F., and Huberman, E. (1988). Perturbation of the cell cycle by adeno-associated virus. Virology *167*, 393–399.

Wistuba, A., Weger, S., Kern, A., and Kleinschmidt, J.A. (1995). Intermediates of adeno-associated virus type 2 assembly: identification of soluble complexes containing Rep and Cap proteins. J. Virol. *69*, 5311–5319.

Wistuba, A., Kern, A., Weger, S., Grimm, D., and Kleinschmidt, J.A. (1997). Subcellular compartmentalization of adeno-associated virus type 2 assembly. J. Virol. *71*, 1341–1352.

Witzgall, R., O'Leary, E., Gessner, R., Ouellette, A.J., and Bonventre, J. V (1993). Kid-1, a putative renal transcription factor: regulation during ontogeny and in response to ischemia and toxic injury. Mol Cell Biol *13*, 1933–1942.

Witzgall, R., O'Leary, E., Leaf, a, Onaldi, D., and Bonventre, J. V (1994). The Krüppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. Proc. Natl. Acad. Sci. U. S. A. *91*, 4514–4518.

Wolf, D., and Goff, S.P. (2007). TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. Cell *131*, 46–57.

Wolf, D., Cammas, F., Losson, R., and Goff, S.P. (2008). Primer binding site-dependent restriction of murine leukemia virus requires HP1 binding by TRIM28. J Virol *82*, 4675–4679.

Wonderling, R.S., and Owens, R.A. (1996). The Rep68 protein of adeno-associated virus type 2 stimulates expression of the platelet-derived growth factor B c-sis proto-oncogene. J Virol *70*, 4783–4786.

Wonderling, R.S., Kyostio, S.R., and Owens, R.A. (1995). A maltose-binding protein/adeno-associated virus Rep68 fusion protein has DNA-RNA helicase and ATPase activities. J Virol *69*, 3542–3548.

Wu, S., Shi, Y.Y., Mulligan, P., Gay, F., Landry, J., Liu, H., Lu, J., Qi, H.H., Wang, W., Nickoloff, J.A., et al. (2007). A YY1-INO80 complex regulates genomic stability through homologous recombination-based repair. Nat. Struct. Mol. Biol. *14*, 1165–1172.

Wu, Z.J., Miller, E., Agbandje-McKenna, M., and Samulski, R.J. (2006). alpha 2,3 and alpha 2,6 N-linked sialic acids facilitate efficient binding and transduction by adeno-associated virus types 1 and 6. J. Virol. *80*, 9093–9103.

Xia, K., Knipe, D.M., and DeLuca, N. a (1996). Role of protein kinase A and the serine-rich region of herpes simplex virus type 1 ICP4 in viral replication. J. Virol. *70*, 1050–1060.

Xie, Q., and Chapman, M.S. (1996). Canine parvovirus capsid structure, analyzed at 2.9 A resolution. J. Mol. Biol. *264*, 497–520.

Xie, Q., Bu, W., Bhatia, S., Hare, J., Somasundaram, T., Azzi, A., and Chapman, M.S. (2002). The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. Proc. Natl. Acad. Sci. U. S. A. *99*, 10405–10410.

Xu, Y., Zhang, H., Nguyen, V.T.M., Angelopoulos, N., Nunes, J., Reid, A., Buluwela, L., Magnani, L., Stebbing, J., and Giamas, G. (2015). LMTK3 Represses Tumor Suppressor-like Genes through Chromatin Remodeling in Breast Cancer. Cell Rep. *12*, 837–849.

Xue, Y., Johnson, J.S., Ornelles, D.A., Lieberman, J., and Engel, D.A. (2005). Adenovirus protein VII functions throughout early phase and interacts with cellular proteins SET and pp32. J Virol *79*, 2474–2483.

Yakobson, B., Koch, T., and Winocour, E. (1987). Replication of adeno-associated virus in

synchronized cells without the addition of a helper virus. J. Virol. *61*, 972–981.

Yakobson, B., Hrynko, T. a, Peak, M.J., and Winocour, E. (1989). Replication of adeno-associated virus in cells irradiated with UV light at 254 nm. J. Virol. *63*, 1023–1030.

Yalkinoglu, A.O., Heilbronn, R., Burkle, A., Schlehofer, J.R., and Zur Hausen, H. (1988). DNA amplification of adeno-associated virus as a response to cellular genotoxic stress. Cancer Res. *48*, 3123–3129.

Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., and Handa, H. (1999). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. Cell *97*, 41–51.

Yamaguchi, Y., Inukai, N., Narita, T., Wada, T., and Handa, H. (2002). Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. Mol. Cell. Biol. *22*, 2918–2927.

Yang, Q., Coseno, M., Gilmartin, G., and Doublie, S. (2011). Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. Structur *19*, 368–377.

Yondola, M. a, and Hearing, P. (2007). The adenovirus E4 ORF3 protein binds and reorganizes the TRIM family member transcriptional intermediary factor 1 alpha. J Virol *81*, 4264–4271.

Yoon, M., Smith, D.H., Ward, P., Medrano, F.J., Aggarwal, A.K., and Linden, R.M. (2001). Amino-terminal domain exchange redirects origin-specific interactions of adeno-associated virus rep78 in vitro. J Virol *75*, 3230–3239.

Yoshizumi, M., Wang, H., Hsieh, C.M., Sibinga, N.E.S., Perrella, M.A., and Lee, M.E. (1997). Down-regulation of the cyclin a promoter by transforming growth factor- β1 is associated with a reduction in phosphorylated activating transcription factor-1 and cyclic AMP-responsive element-binding protein. J. Biol. Chem. *272*, 22259–22264.

Young Jr., S.M., and Samulski, R.J. (2001). Adeno-associated virus (AAV) site-specific recombination does not require a Rep-dependent origin of replication within the AAV terminal repeat. Proc Natl Acad Sci U S A *98*, 13525–13530.

Zaiss, A.K., Cotter, M.J., White, L.R., Clark, S.A., Wong, N.C., Holers, V.M., Bartlett, J.S., and Muruve, D.A. (2008). Complement is an essential component of the immune response to adeno-associated virus vectors. J. Virol. *82*, 2727–2740.

Zarate-Perez, F., Bardelli, M., Burgner, J.W., Villamil-Jarauta, M., Das, K., Kekilli, D., Mansilla-Soto, J., Linden, R.M., and Escalante, C.R. (2012). The interdomain linker of AAV-2 Rep68 is an integral part of its oligomerization domain: Role of a conserved SF3 helicase residue in oligomerization. PLoS Pathog. *8*, 1–12.

Zarate-Perez, F., Mansilla-Soto, J., Bardelli, M., Burgner, J.W., Villamil-Jarauta, M., Kekilli, D., Samso, M., Linden, R.M., and Escalante, C.R. (2013). Oligomeric properties of adeno-associated virus Rep68 reflect its multifunctionality. J. Virol. *87*, 1232–1241.

Zeltner, N., Kohlbrenner, E., Clement, N., Weber, T., and Linden, R.M. (2010). Near-perfect infectivity of wild-type AAV as benchmark for infectivity of recombinant AAV vectors. Gene Ther *17*, 872–879.

Zeng, L., Yap, K.L., Ivanov, A. V, Wang, X., Mujtaba, S., Plotnikova, O., Rauscher 3rd, F.J., and Zhou, M.M. (2008). Structural insights into human KAP1 PHD finger-bromodomain and its role in gene silencing. Nat Struct Mol Biol *15*, 626–633.

Zhang, L., Zhu, C., Guo, Y., Wei, F., Lu, J., Qin, J., Banerjee, S., Wang, J., Shang, H.,

Verma, S.C., et al. (2014). Inhibition of KAP1 enhances hypoxia-induced Kaposi's sarcoma-associated herpesvirus reactivation through RBP-Jkappa. J Virol *88*, 6873–6884.

Zincarelli, C., Soltys, S., Rengo, G., and Rabinowitz, J.E. (2008). Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. Mol. Ther. *16*, 1073–1080.

Zincarelli, C., Soltys, S., Rengo, G., Koch, W.J., and Rabinowitz, J.E. (2010). Comparative Cardiac Gene Delivery of Adeno-Associated Virus Serotypes 1–9 reveals that AAV6 Mediates the Most Efficient Transduction in Mouse Heart. Clin. Transl. Sci. *3*, 81–89.

Ziv, Y., Bielopolski, D., Galanty, Y., Lukas, C., Taya, Y., Schultz, D.C., Lukas, J., Bekker-Jensen, S., Bartek, J., and Shiloh, Y. (2006). Chromatin relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. Nat Cell Biol *8*, 870–876.

APPENDIX 1: SUPPLEMENTARY INFORMATION

Table 1. Primers

Name	Sequence (5'-> 3')	Target
	Cloning Primers	
SS13	Fw - GGCGGAGGAGGCGGAGGAGGCGGAGGAGGC ATGGAGCTGGTCGGGTGGCT	Rep40 and Rep52 from pND229/234 and pND230/233
SS14	Rv – CCAAGGGCCCTCAGAGAGAGTGTCCTC GAGCCAAT	Rep40 and Rep68 from pND229/234 and pND226/232
SS16	Rv – CCAAGGGCCCTTATTGTTCAAAGATGCAGTC ATCCAAATC	Rep52 and Rep78 from pND230/233 and pND227/231
SS22	Rv – GCCTCCTCCGCCTCCTCCGCCTCCTCCGCCctcg agcttctctgcgcttctcaggg	BirA* from pMycBirA*
SS23	Fw – CCAAGGGCCCgctagccaccatggaacaaaaactcat	BirA* from pMycBirA*
SS40	Fw - GGCGGAGGAGGCGGAGGAGGCGGAGGAGGCA TGCCGGGGTTTTACGAGATTGTGA	Rep68 and Rep78 from pND226/232 and pND227/231
SS52	Fw – ccaagaattcgggagccaccatggattacaaggatgacgatgacaa gggaggaggattggtgataccccccggactgagcgaggaagag	RDBP from IMAGE clone
SS53	Rv - ccaactcgagttccctagaagccatccacaaggttttccttgtagacgtc atcactgtagac	RDBP from IMAGE clone
SS54	Fw – ccaagaattcgctagccaccatggattacaaggatgacgatgacaag ggaggaggaaagattgaggaggtgaagagcactacgaagacgca gcgc	RUVBL1 from IMAGE clone
SS55	Rv – ccaactcgagccatctcacttcatgtacttatcctgctggtcagccagga tttt	RUVBL1 from IMAGE clone
SS57	Rv - ccaactcgagagaattattttgttgatcgtctcgtgggctccgatgct	TCERG1 from IMAGE clone
SS59	Fw - ccaactcgagagaattacttgtcatcgtcatccttgtaatctcctcctcctt tgttgatcgtctcgtgggctccgatgctgtgggaggtggg	TCERG1 from IMAGE clone

SS62	Fw - ccaaTCTAGAGCCGTGAGCAAGGGCGAGGAGCTGTTC	GFP from pND104
SS63	Rv - ccaaCTCGAGTGAATTATTACTTGTACAGCTCGTCCATGCCGAGAG	GFP from pND104
SS81	Fw – CCAACTCGAGGCGCGCCACCATGGACTACAAAGACGATG	FLAG-KAP1 from pKerppola
SS82	Rv – RV primer for TRIM28 containing XhoI site	FLAG-KAP1 from pKerppola
SS93	Fw – ccaaTCTAGAGCCtccgacagcgagaagctcaacctggac	PP1 from IMAGE clone
SS94	Rv – ccaaCTCGAGTGAActatttcttggctttggcggaattgcgggg	PP1 from IMAGE clone
SS100	Fw – CTTTCCCTTCAACGACTGTGTCGACgcgATGGTGATCTGGTGGGAGGAGGGG	Mutagenesis on mini-pDG and Rep constructs
SS101	Rv – CCCCTCCTCCCACCAGATCACCATcgcGTCGACACAGTCGTTGAAGGGAAAG	Mutagenesis on mini-pDG and Rep constructs
SS102	Fw – ccaaTCTAGAGTGAGCAAGGGCGAGGAGCTGTTCAC	GFP from pSS015
SS103	Rv – ccaaAGATCTTTACTTGTACAGCTCGTCCATGCCGAGAGTG	GFP from pSS015
SS107	FW – ccaaTCTAGAGCCGAGCTGGTCGGGTGGCTCGTG	Rep52 and Rep52 from pSS024 and pSS025
SS108	RV – ccaaTCTAGACCCGGGTTATTGTTCAAAGATGCAGTCATCC	Rep52 and Rep78 from pSS025 and pSS027
SS109	FW – ccaaTCTAGAGCCCCGGGGTTTTACGAGATTGTGATTAAGGTCCC	Rep68 and Rep78 from pSS026 and pSS027
SS110	RV – ccaaTCTAGACCCGGGTCAGAGAGAGTGTCCTCGAG	Rep40 and Rep68 from pSS024 and pSS026
SS117	FW – ccaaTCTAGAggaggcGCGGCCTCCGCGGCGGCAGCC	KAP1 from pKerppola
SS118	RV – ccaaAGATCTTTAGGGGCCATCACCAGGGCCACCAGACAGC	KAP1 from pKerppola

SS119	FW – ccaaGTCGACAAGATGGTGATCTGG	Rep52 and Rep78 from pSS025 and pSS027
SS123	RV – ccaaCCCGGGTTAGTACCTGTCTGCGTAGTTGAT CGAAGCTTCC	Rep52 and Rep78 from pSS025 and pSS027
SS124	RV – ccaaCCCGGGTTAGATATTTGAATTCTGATTCATT CTCTCGCATTG	Rep52 and Rep78 from pSS025 and pSS027
SS126	RV – ccaaCCCGGGTTATTGAGATTCTGACACGGGAAA GCACTCTAAAC	Rep52 and Rep78 from pSS025 and pSS027
SS129	FW – ccaaCTCGAGgccaccatgGCGGCCTCCGCGGCGG CAG	KAP1 from pKerppola
qPCR primers		
Cap1	TTCTCAGATGCTGCGTACCGGAAA	WT AAV quantification
Cap2	TCTGCCATTGAGGTGGTACTTGGT	WT AAV quantification
Ad5 100K_FW	TCATTACCCAGGGCCACATT	Ad5 quantification
Ad5 100k_RV	CCTCGTCCAAAACCTCCTCT	Ad5 quantification
Cyclo_FW	TGCTGGACCCAACACAAATG	Housekeeping control
Cyclo_RV	TGCCATCCAACCACTCAGTCT	Housekeeping control
GAPDH_FW	CACCGTCAAGGCTGAGAACG	ChIP-qPCR
GAPDH_RV	ATACCCAAGGGAGCCACACC	ChIP-qPCR
ZNF180_FW	TGATGCACAATAAGTCGAGCA	ChIP-qPCR
ZNF180_RV	TGCAGTCAATGTGGGAAGTC	ChIP-qPCR
ZNF274_FW	GGAGAAATCCCATGAGGGTAA	ChIP-qPCR
ZNF274_RV	GGCTTTTGTGAGAATGTTTTCC	ChIP-qPCR
p5_FW	CTGTATTAGAGGTCACGTGAGTG	ChIP-qPCR
p5_RV	TCAAACCTCCCGCTTCAAA	ChIP-qPCR
Rep5'_FW	CCGAGAAGGAATGGGAGTT	ChIP-qPCR

Rep5'_RV	CCATTCCGTCAGAAAGTCG	ChIP-qPCR
Rep middle_FW	GCCTTGGACAATGCGGGAAAGATT	ChIP-qPCR
Rep middle_RV	TGTCGACACAGTCGTTGAAGGGAA	ChIP-qPCR
Rep3'_FW	TTCCCGTGTCAGAATCTCAA	ChIP-qPCR
Rep3'_RV	CCAAATCCACATTGACCAGA	ChIP-qPCR
Cap5'_FW	GACAGTGGTGGAAGCTCAAA	ChIP-qPCR
Cap5'_RV	TTGTACCCAGGAAGCACAAG	ChIP-qPCR
Cap middle_FW	TTCTCAGATGCTGCGTACCGGAAA	ChIP-qPCR
Cap middle_RV	TCTGCCATTGAGGTGGTACTTGGT	ChIP-qPCR
Cap3'_FW	GTCAGCGTGGAGATCGAGT	ChIP-qPCR
Cap3'_RV	AGGCTCTGAATACACGCCAT	ChIP-qPCR
MB85_p5fw	AACAAGGTGGTGGATGAGT	taqman qPCR on AAV cDNA
MB86_p5rv	CGTTTACGCTCCGTGAGATT	taqman qPCR on AAV cDNA
MB87_p19fw	TCACCAAGCAGGAAGTCAAAG	taqman qPCR on AAV cDNA
MB88_p19rv	CCCGTTTGGGCTCACTTATATC	taqman qPCR on AAV cDNA
MB89_p40fw	GGAAGCAAGGCTCAGAGAAA	taqman qPCR on AAV cDNA
MB90_p40rv	CCTCTCTGGAGGTTGGTAGATA	taqman qPCR on AAV cDNA
Primers for mutagenesis		
SS100_FW	CTTTCCCTTCAACGACTGTGTCGACgcgATGGTG ATCTGGTGGGAGGAGGGG	K372A mutagenesis on Rep52 and Rep78
SS101_RV	CCCCTCCTCCCACCAGATCACCATcgcGTCGACA CAGTCGTTGAAGGGAAAG	K372A mutagenesis on Rep52 and Rep78
qPCR probes		
Probe p5	FAM-ACTGTTCCATATTAGTCCACGCCCAC-TAM	FAM-TAMRA probe to b used with MB85 and MB86
Probe p19	FAM-ACGTGGTTGAGGTGGAGCATGAAT-TAM	FAM-TAMRA probe to b used with MB87 and MB88
Probe p40	FAM-AGGAAATCAGGACAACCAATCCCGT-TAM	FAM-TAMRA probe to b used with MB89 and MB90

Table 2. Plasmids

Plasmid	Content	Cloning strategy	Resistance
Rep			
pND229	CMV-Rep40	Cloned by Nathalie Dutheil	Kana
pND230	CMV-Rep52	Cloned by Nathalie Dutheil	Kana
pND226	CMV-Rep68 Y156 M225G	Cloned by Nathalie Dutheil	Kana
pND227	CMV-Rep78 Y156 M225G	Cloned by Nathalie Dutheil	Kana
pND234	CMV-Rep40 K340H	Cloned by Nathalie Dutheil	Kana
pND233	CMV-Rep52 K340H	Cloned by Nathalie Dutheil	Kana
pND232	CMV-Rep68 Y156 M225G K340H	Cloned by Nathalie Dutheil	Kana
pND231	CMV-Rep78 Y156 M225G K340H	Cloned by Nathalie Dutheil	Kana
pSS003	CMV-BirA*-Rep40	Fragment amplified from mycBirA* and pND229 using SS23/22 (BirA*), SS13/14 (Rep40), and SS23/14 (BirA*-Rep40); cloned into pcDNA3.1+ vector using ApaI	Amp
pSS004	CMV-BirA*-Rep40 K340H	Fragment amplified from mycBirA* and pND234 using SS23/22 (BirA*), SS13/14 (Rep40), and SS23/14 (BirA*-Rep40); cloned into pcDNA3.1+ vector using ApaI	Amp
pSS005	CMV-BirA*-Rep52	Fragment amplified from mycBirA* and pND230 using SS23/22 (BirA*), SS13/16 (Rep52), and SS23/16 (BirA*-Rep52); cloned into pcDNA3.1+ vector using ApaI	Amp
pSS006	CMV-BirA*-Rep52 K340H	Fragment amplified from mycBirA* and pND233 using SS23/22 (BirA*), SS13/16 (Rep52), and SS23/16 (BirA*-Rep52); cloned into pcDNA3.1+ vector using ApaI	Amp
pSS009	CMV-BirA*-Rep68 Y156 M225G	Fragment amplified from mycBirA* and pND226 using SS23/22 (BirA*), SS40/14 (Rep68), and SS23/14 (BirA*-Rep68); cloned into pcDNA3.1+ vector using ApaI	Amp
pSS010	CMV-BirA*-Rep68 Y156 M225G K340H	Fragment amplified from mycBirA* and pND232 using SS23/22 (BirA*), SS40/14 (Rep68), and SS23/14 (BirA*-Rep68); cloned into pcDNA3.1+ vector using ApaI	Amp
pSS011	CMV-BirA*-Rep78 Y156 M225G	Fragment amplified from mycBirA* and pND227 using SS23/22 (BirA*), SS40/16 (Rep78), and SS23/16 (BirA*-Rep78); cloned into pcDNA3.1+ vector using ApaI	Amp
pSS012	CMV-BirA*-Rep78 Y156 M225G K340H	Fragment amplified from mycBirA* and pND231 using SS23/22 (BirA*), SS40/16 (Rep78), and SS23/16 (BirA*-Rep78); cloned into pcDNA3.1+ vector using ApaI	Amp

pSS024	CMV-T7-Rep40	Rep amplified from pND229 using SS85/87; cloned into T7-tagged vector using XbaI/SmaI	Amp
pSS025	CMV-T7-Rep52	Rep amplified from pND230 using SS85/88; cloned into T7-tagged vector using XbaI/SmaI	Amp
pSS026	CMV-T7-Rep68 Y156 M225G	Rep amplified from pND226 using SS86/87; cloned into T7-tagged vector using XbaI/SmaI	Amp
pSS027	CMV-T7-Rep78 Y156 M225G	Rep amplified from pND227 using SS86/88; cloned into T7-tagged vector using XbaI/SmaI	Amp
pSS033	CMV-T7-Rep40 K340H	Rep amplified from pND234 using SS85/87; cloned into T7-tagged vector using XbaI/SmaI	Amp
pSS034	CMV-T7-Rep52 K340H	Rep amplified from pND233 using SS85/88; cloned into T7-tagged vector using XbaI/SmaI	Amp
pSS036	CMV-T7-Rep68 Y156 M225G K340H	Rep amplified from pND232 using SS86/87; cloned into T7-tagged vector using XbaI/SmaI	Amp
pSS037	CMV-T7-Rep78 Y156 M225G K340H	Rep amplified from pND231 using SS86/88; cloned into T7-tagged vector using XbaI/SmaI	Amp
pSS040	CMV-T7-Rep52 K372A	Mutagenesis on pSS025 using SS100/101	Amp
pSS041	CMV-T7-Rep78 Y156 M225G K372A	Mutagenesis on pSS027 using SS100/101	Amp
pSS042	CMV-T7-Rep52 K340H K372A	Mutagenesis on pSS034 using SS100/101	Amp
pSS043	CMV-T7-Rep78 Y156 M225G K340H K372A	Mutagenesis on pSS037 using SS100/101	Amp
pSS047	CMV-Rep52 K372A	Rep amplified from pSS025 and cloned into pcDNA3.1+ using SmaI	Amp
pSS048	CMV-Rep78 Y156 M225G K372A	Rep amplified from pSS027 and cloned into pcDNA3.1+ using SmaI	Amp
pSS058	CMV-T7-Rep52Δ87	Truncated Rep amplified from pSS025 using SS119/123 and cloned in T7-tagged vector using HincII/SmaI	Amp
pSS059	CMV-T7-Rep52Δ62	Truncated Rep amplified from pSS025 using SS119/124 and cloned in T7-tagged vector using HincII/SmaI	Amp

pSS060	CMV-T7-Rep52Δ43	Truncated Rep amplified from pSS025 using SS119/126 and cloned in T7-tagged vector using HincII/SmaI	Amp
pSS061	CMV-T7-Rep78 Y156 M225G Δ91	Truncated Rep amplified from pSS027 using SS119/123 and cloned in T7-tagged vector using HincII/SmaI	Amp
pSS062	CMV-T7-Rep78 Y156 M225G Δ62	Truncated Rep amplified from pSS027 using SS119/124 and cloned in T7-tagged vector using HincII/SmaI	Amp
pSS063	CMV-T7-Rep78 Y156 M225G Δ43	Truncated Rep amplified from pSS027 usingSS119/126 and cloned in T7-tagged vector using HincII/SmaI	Amp
KAP1			
pKerppola	CMV-FLAG-KAP1 (full length KAP1)	Obtained from Helen Rowe	Kana
pSS037	CMVΔ5-FLAG-KAP1	FLAG-KAP1 amplified from pKerppola using SS81/82 and cloned into pWISP12-96 vector using XhoI	Amp
pSS038	CMVΔ6-FLAG-KAP1	FLAG-KAP1 amplified from pKerppola using SS81/82 and cloned into pWISP12-97 vector using XhoI	Amp
pSS064	CMV-T7-KAP1	KAP1 amplified from pKerppola using SS117/118 and cloned into T7 vector using XbaI/BglII digest on fragment and XbaI/BamHI digest on vector	Amp
pSS074	CMV-KAP1	KAP1 amplified from pKerppola using SS129/SS82 and cloned into pcDNA3.1+ using XhoI	Amp
Others			
pSS014	FLAG-TCERG1	FLAG-TCERG1 amplified from IMAGE clone using SS59/57 and cloned into pCDNA3.1+ using XhoI	Amp
pIN001	FLAG-RDBP	FLAG-RDBP amplified from IMAGE clone using SS52/53 and cloned into pCDNA3.1+ using EcoRI/XhoI	Amp
pIN002	FLAG-RUVBL1	FLAG-RUVBL1 amplified from IMAGE clone using SS54/55 and cloned into pCDNA3.1+ using NotI/XhoI	Amp
pSS015	FLAG-GFP	GFP amplified from pND104 using SS62/63 and cloned into pKerppola backbone containing FLAG tag using XbaI/XhoI	Kana
pSS028	FLAG-PP1α	PP1 amplified from IMAGE clone using SS93/94 and cloned into FLAG vector using XbaI/XhoI	Amp

pSS044	T7-GFP	GFP amplified from pSS015 using SS102/103 and cloned into T7 vector using XbaI/BglII digest on fragment and XbaI/BamHI digest on vector	Amp
pEXN-CPSF6	CMV-CPSF6	Obtained from Greg Towers	Amp
pC1-NIPP1	GFP-NIPP1	Obtained from Angus Lammond via Addgene	Amp
AAV/rAAV/Helpers			
pAV2	Infectious AAV plasmid		Amp
pMB32	pMB2 based, Rep mutation to K340H	Cloned by Martino Bardelli	Amp
pSS078	pAV2 based, Rep mutation to K372A	Mutagenesis on mini-pDG using SS100/101 then cloned into pAV2 using SfiI/HindIII	Amp
pDG	Rep and Cap, and Ad helper functions		Amp
Mini-pDG	Rep and Cap only	Cloned by Els Henckaerts	Amp
pTRUF11	CAG-GFP-pk-NEO between ITRs		Amp

Table 3. siRNA

siRNA	Target Sequence	Source
siTCERG1	GGAGUUGCACAAGAUAGUU	Sune et al. 2012
siRUVB1	UAAAGGAGACCCAAGGAAGU	Dutta et al. 2008
siRDBP	GGCAUUGCUGGCUCUGAAG	Li et al. 2004
siKAP1 (Smartpool)	siKAP1.1 - GAAAUGUGAGCGUGUACUG siKAP1.2 - GCGAUCUGGUUAUGUGCAA siKAP1.3 - AGACAGCACUGGCGUGGUG siKAP1.4 - GAACGAGGCCUUCGGUGAC	Dharmacon (L-005046-00-0005)
siCHD3 (Smartpool)	GAAUAUCCCUGAAUACGAA CCAGAAUGAUGCUCAAUUU CAUAAGAGGCGGAGUAAGA CGUAUGAGCUGAUCACCAU	Dharmacon (L-023015-00-0005)
siSETDB1 (Smartpool)	AAGAUGGGCUUUCAUGUUA GAACUAAGACUUGGCACAA GGGAUCAACCAGACAUAUA GAGCGCACCUGUUCGUAAG	Dharmacon (L-020070-00-0005)

Table 4. List of antibodies

Antibody	Source	Conditions	Application
α-HSP90	Santa-Cruz Sc-69703	1:10,000 5% milk in PBST	WB
α-KAP1	Chemicon MAB3662	1:1,000 5% milk in PBST	WB
α-p-KAP1-S824	Bethyl A300-767A	1:2,000 2.5% BSA inPBST	WB IF
α-VP	ARP 03-61058	1:500 5% milk in PBST	WB
α-Rep	Progen 61069	1:100 5% milk in PBST	WB
α-dsRed	Clontech 632496	1:2,000 5% milk in PBST	WB
α-CHD3	Bethyl A301-219A	1:4,000 5% milk in PBST	WB
α-SETDB1	Abcam Ab12317	1:1,000 5% milk in PBST	WB
Rabbit α-FLAG	Sigma F7425	4µg/mL lysate	IP
Mouse α-FLAG	Sigma F1804	1:1,000 5% milk in PBST	WB
α-T7	Merck Millipore 69522-3	WB: 1:10,000 PBST IP: 1µl/500µl lysate	WB IP
α-p-Chk2	NEB 2661S	1:1000 2.5% BSA in PBST	WB
α-GFP	Roche 11814460001	1:5,000 5% milk in PBST	WB
α-rabbit IgG	BioRad	1:10,000 5% milk in PBST	WB
α-mouse IgG	BioRad	1:10,000 5% milk in PBST	WB
α-TCERG1	Bethyl A300-360A	1:5,000 5% milk in PBST	WB
α-RUVB1	Serum obtained from Dr Anindya Dutta	1:1,000 5% milk in PBST	WB
α-RDBP	Serum obtained from Dr Rong Li	1:200 5% milk in PBST	WB
α-CPSF6	Abcam 99347	WB: 1:4,000 5% milk in PBST IP: 2µg/mL lysate	WB IP
α-KAP1	Abcam ab10483	10mg/IP	ChIP
α-H3K9me3	Abcam ab8893	4mg/IP	ChIP
IgG	Abcam ab37415	5mg/IP	ChIP

APPENDIX 2: BioID PULL-DOWN DATA

Table 5. Unique peptides identified by BioID using bait proteins BirA*-Rep52 and BirA*-Rep52-K340H

Identified Proteins	Accession Number	BirA*-Rep52	BirA*-Rep52-K340H	Mock Tfxn.
Cationic trypsin OS=Bos taurus	P00760	10	9	4
HNRNPK	P61978	13	18	0
NASP	P49321	8	12	0
BirA OS=Escherichia coli	P06709	19	24	0
NCL	P19338	3	13	0
REP=Adeno-associated virus 2	P03132	15	20	0
SF3B1	O75533	5	11	0
FLNA	P21333	2	14	0
RBM26	Q5T8P6	4	7	0
SF3B2	Q13435	3	11	0
NUDC	Q9Y266	2	11	0
CCT8	P50990	4	8	0
SNW1	Q13573	4	9	0
SF3A1	Q15459	4	9	0
TCERG1	O14776	2	9	0
TMPO	P42167	4	8	0
SF1	Q15637	3	7	0
YLMP1	P49750	0	10	0
SFPQ	P23246	5	6	0
HNRNPU	Q00839	3	8	0
TAGLN2	P37802	1	9	0
TOX4	O94842	5	5	0
DACH1	Q9UI36	5	5	0
TRIM28	Q13263	4	4	0
HSPA1A	P08107	4	5	0
MRE11A	P49959	0	8	0
CPSF6	Q16630	3	6	0
TCOF1	Q13428	3	4	0
RUVBL1	Q9Y265	4	5	0
BCLAF1	Q9NYF8	3	5	0
DNAJC8	O75937	3	4	0
THRAP3	Q9Y2W1	2	5	0
SERBP1	Q8NC51	1	7	0
HNRNPA1	P09651	0	8	0

CFDP1	Q9UEE9	2	5	0
U2SURP	O15042	0	7	0
FGA	P02671	3	4	0
PCNP	Q8WW12	2	5	0
HSP90AB1	P08238	0	7	0
WDR70	Q9NW82	0	7	0
CNN3	Q15417	1	5	0
SSB	P05455	0	6	0
HCHC1	P51610	0	6	0
USP15	Q9Y4E8	0	5	0
UBB	P0CG47	2	3	0
HIST1H4A	P62805	3	2	0
PDLIM1	O00151	0	5	0
PPIL4	Q8WUA2	0	6	0
Avidin OS=Gallus gallus	P02701	3	8	0
ZRANB2	O95218	2	3	0
MDC1	Q14676	1	4	0
SPDL1	Q96EA4	0	5	0
PHAX	Q9H814	0	5	0
BRD4	O60885	0	3	0
NONO	Q15233	0	4	0
HIST1H2AB	P04908	1	3	0
HNRNPD	Q14103	1	3	0
VBP1	P61758	0	4	0
SUGT1	Q9Y2Z0	0	4	0
SART1	O43290	0	4	0
ACIN1	Q9UKV3	0	4	0
NUP50	Q9UKX7	0	4	0
NPM1	P06748	1	2	0
FAM50A	Q14320	0	3	0
ORC2	Q13416	0	3	0
Trypsin-1 OS=Homo sapiens	P07477	1	1	1
GTF2A1	P52655	1	2	0
CRKL	P46109	1	2	0
NUCKS1	Q9H1E3	1	2	0
HNRNPA2B1	P22626	0	3	0
SPAG7	O75391	0	3	0
API5	Q9BZZ5	0	3	0
ENO1	P06733	0	3	0
NUDT5	Q9UKK9	0	3	0
KHSRP	Q92945	0	3	0
PPP1R10	Q96QC0	0	3	0
CDC5L	Q99459	0	3	0

GMEB2	Q9UKD1	0	2	0
SNRPA1	P09661	0	3	0
RPRD2	Q5VT52	0	2	0
PUS7	Q96PZ0	0	3	0
CHERP	Q8IWX8	0	3	0
EEF1A1	P68104	1	1	0
HIST1H2BK	O60814	0	2	0
SNRNP200	O75643	0	2	0
WAC	Q9BTA9	0	2	0
RNF113A	O15541	0	2	0
C3	P01024	1	1	0
ARID3B	Q8IVW6	1	1	0
MAP4	P27816	0	2	0
RBM10	P98175	0	2	0
NACA	Q13765	0	2	0
SAMHD1	Q9Y3Z3	0	2	0
GTF2E2	P29084	0	2	0
PPM1G	O15355	0	2	0
DNAJB1	P25685	0	2	0
IK	Q13123	0	2	0
HNRNPL	P14866	0	2	0
SAFB2	Q14151	0	2	0
ZMYND8	Q9ULU4	0	2	0
PKM	P14618	0	2	0
TLE1	Q04724	0	2	0
RDBP	P18615	0	2	0
SMEK1	Q6IN85	0	2	0
SPEN	Q96T58	0	2	0
CCDC174	Q6PII3	0	2	0
RFC4	P35249	0	2	0
NUMA1	Q14980	0	2	0
ZNF318	Q5VUA4	0	2	0
TPX2	Q9ULW0	1	1	0
KRIT1	O00522	1	0	0
EIF4B	P23588	0	1	0
RBM17	Q96I25	0	1	0
MKI67	P46013	0	1	0
SUMO3	P55854	1	0	0
AKAP8	O43823	0	1	0
PC	P11498	0	1	0
PIP	P12273	0	1	0
TAF1	P21675	0	1	0
TKT	P29401	0	1	0

FUS	P35637	0	1	0
CBX5	P45973	0	1	0
CTCF	P49711	0	1	0
PAPOLA	P51003	0	1	0
MNAT1	P51948	0	1	0
PSME3	P61289	0	1	0
SNRPE	P62304	0	1	0
CHAF1A	Q13111	0	1	0
RIF1	Q5UIP0	0	1	0
CIAPIN1	Q6FI81	0	1	0
RAI1	Q7Z5J4	0	1	0
WAPAL	Q7Z5K2	0	1	0
C15orf38	Q7Z6K5	0	1	0
BOD1L1	Q8NFC6	0	1	0
DGCR8	Q8WYQ5	0	1	0
CCDC104	Q96G28	0	1	0
FTO	Q9C0B1	0	1	0
LUC7L	Q9NQ29	0	1	0
COPS7A	Q9UBW8	0	1	0
FAM208A	Q9UK61	0	1	0
LARP1	Q6PKG0	0	1	0
SMAP	O00193	0	1	0
NOL11	Q9H8H0	0	1	0
CD2AP	Q9Y5K6	0	1	0
EIF2A	Q9BY44	0	1	0
MATR3	P43243	0	1	0
NSRP1	Q9H0G5	0	1	0
EIF5A	P63241	0	1	0
EIF5	P55010	0	1	0
PYGL	P06737	0	1	0
DLGAP5	Q15398	0	1	0
HNRNPH1	P31943	0	1	0
PRCC	Q92733	0	1	0
USP5	P45974	0	1	0
ARID1A	O14497	0	1	0
ILKAP	Q9H0C8	0	1	0
CCT6A	P40227	0	1	0
NCOR1	O75376	0	1	0
MED1	Q15648	0	1	0
ANP32E	Q9BTT0	0	1	0
DUT	P33316	0	1	0
RBBP6	Q7Z6E9	0	1	0
HNRNPR	O43390	0	1	0

GCFC2	P16383	0	1	0
CRK	P46108	0	1	0
RBM12	Q9NTZ6	0	0	0
Serum albumin OS=Bos taurus	P02769	0	0	1

Table 6. Unique peptides identified by BioID using bait proteins BirA*-Rep40 and BirA*-Rep40-K340H

Identified Proteins	Accession number	BirA*-Rep40	BirA*-Rep40-K340H
BirA=Escherichia coli	P06709	11	19
REP=Adeno-associated virus 2	P03132	7	9
TAGLN2	P37802	4	7
HNRNPK	P61978	5	8
SF1	Q15637	2	6
SNW1	Q13573	2	5
PDLIM1	O00151	1	5
NASP	P49321	2	4
NPM1	P06748	3	2
GTF2E2	P29084	1	2
TMPO	P42167	1	2
PUS7	Q96PZ0	1	2
HNRNPU	Q00839	1	3
SF3A1	A2VDN6		5
FLNA	P21333		5
CHMP5	Q5RBR3		4
DNAJB1	P25685		4
DUT	P33316		3
PCMT1	P22061		2
SPAG7	O75391	1	1
SLC4A1AP	Q9BWU0		4
TPX2	Q9ULW0		4
HDGF	P51858		2
TRIM28	Q13263		3
ENO1	P06733	1	2
EIF5A	P10160	1	2
CLIC1	O00299		3
TCERG1	O14776		3
TOX4	O94842		3
YLPM1	P49750		3
USP15	Q2HJE4		3
SPDL1	Q96EA4		3
SF3B2	Q13435		3
CWC15	Q2KJD3	1	1
HSPA1A	P08107		2
RDBP	P18615		2
SFPQ	P23246		2
CPSF6	Q0P5D2		2

CTTN	Q14247		2
MDC1	Q14676		2
GPKOW	Q92917		2
MCCC1	Q96RQ3		2
PHAX	Q9H814		2
DNAJC8	O75937		2
HNRNPA1	A5A6H4		2
RUVBL1	P60122		2
PPIL4	Q8WUA2		2
KRT4	P19013	1	
SMNDC1	O75940		2
SAP30B	Q02614		2
NUP50	Q9UKX7		2
CHAMP1	Q96JM3		2
ZC3H4	Q9UPT8		2
CRKL	P46109		2
STMN1	A9YWH3		1
SF3B1	O57683		1
ZC3H14	Q4R6F6		1
SERBP1	Q6AXS5		1
C15orf38	Q7Z6K5		1
FAM192A	Q91WE2		1
PIPSL	A2A3N6		1
EFHD2	A5D7A0		1
ALYREF	B5FXN8		1
SMAP	O00193		1
DFFA	O00273		1
PDCD5	O14737		1
PPMIG	O15355		1
RNF113A	O15541		1
NUDC	O35685		1
API5	O35841		1
TXNL1	O43396		1
HTATSF1	O43719		1
EDF1	O60869		1
ZC3H11A	O75152		1
CFDP1	O88271		1
NCL	P08199		1
PGAM1	P18669		1
ARHGDIA	P19803		1
S100A7	P31151		1
PTPN11	P35235		1
FUS	P35637		1

FASN	P49327		1
MRE11A	P49959		1
CCT8	P50990		1
ADAR	P55265		1
VBP1	P61758		1
SET	Q01105		1
WDR70	Q0VA16		1
PPP1R8	Q12972		1
IK	Q13123		1
ORC2	Q13416		1
NT5DC1	Q2TBU5		1
PCNP	Q32PF3		1
HN1	Q3T0T5		1
THRAP3	Q569Z6		1
NUDT5	Q5RCY2		1
CHMP4B	Q5XGW6		1
PRCC	Q92733		1
NANS	Q9NR45		1
SIX4	Q9UIU6		1
KRT16	P08779	1	
SNRPA1	P09661		1
SNRNP40	Q5RF51		1
GTF2E1 PE=2 SV=1	A6QLI8		1
TAF9	Q16594		1
HNRPDL	O14979		1
BCLAF1	Q9NYF8		1
HNRNPL	P14866		1
HNRNPL	P14866		1

Table 7. Unique peptides identified by BioID using bait proteins BirA*-Rep68-K340H and BirA*-Rep78-K340H

Gene Name	Accession number	BirA*-Rep68-K340H	BirA*-Rep78-K340H
TAGLN2	P37802	23	18
FLNA	P21333	21	10
Trypsin-1 OS=Homo sapiens	P07477	15	15
PCNP	Q8WW12	14	9
NCL	P19338	15	4
Serum albumin OS=Homo sapiens	P02768	9	14
SERBP1	Q8NC51	12	4
NASP	P49321	9	5
ACTB	P60709	6	4
HNRNPK	P61978	11	4
DSG1	Q02413	4	4
HSP90AB1	P08238	9	2
DNAJC8	O75937	4	2
HN1L	Q9H910	5	3
FUS	P35637	5	3
CNN3	Q15417	6	4
CWC15	Q9P013	2	3
VIM	P08670	4	
SF1	Q15637	4	2
NPM1	P06748	5	2
SET	Q01105	4	1
AMY1A	P04745	3	2
ANXA2	P07355	2	2
GAPDH	P04406	3	1
SERPINB3	P29508	1	2
PDLIM1	O00151	5	2
DSP	P15924	1	2
DCD	P81605	3	3
PRDX1	Q06830	3	1
SMAP	O00193	3	4
TMPO	P42166	5	2
TRIM28	Q13263	3	1
AZGP1	P25311	2	3
UBB	P0CG47	3	1
CTTN	Q14247	2	1
S100A9	P06702		4
HMGB1	P09429	5	
S100A8	P05109	1	3

HSPA8	P11142	2	1
SAP30BP	Q9UHR5	1	2
C1orf52	Q8N6N3	2	2
TAF9	Q16594	4	2
HDGF	P51858	3	1
KIAA1704	Q8IXQ4	2	2
SPAG7	O75391	3	2
EDF1	O60869	3	2
TOX4	O94842	2	2
NUCKS1	Q9H1E3	4	2
CFDP1	Q9UEE9	2	2
NACA	Q13765	4	1
GTF2E2	P29084	2	1
C15orf38	Q7Z6K5	5	
GGCT	O75223	1	1
FUBP1	Q96AE4	4	
HNRNPD	Q14103	4	2
FAM192A	Q9GZU8	2	2
SNW1	Q13573	1	2
KHSRP	Q92945	3	1
PDAP1	Q13442	3	1
TUBA1B	P68363	2	
HSP90AA1	P07900	8	
JUP	P14923	2	
RDBP	P18615	4	
EPRS	P07814	2	2
DSC1	Q08554	1	2
SERPINB12	Q96P63	1	1
HBB	P68871	1	2
HIST1H2AB	P04908	2	1
YWHAE	P62258	2	
STMN1	P16949	2	
TCOF1	Q13428	3	
CIAPIN1	Q6FI81	3	
PIP	P12273	1	1
ARG1	P05089	1	1
PHAX	Q9H814	2	1
ZRANB2	O95218	2	1
EIF4H	Q15056	2	1
HSPA5	P11021	2	
FABP5	Q01469	1	
KPRP	Q5T749		1
CASP14	P31944	1	

EEF1A1	P68104	1	
NUDT5	Q9UKK9	3	
HIST1H4A	P62805	2	
PRSS2	P07478	2	2
CALD1	Q05682	1	1
ARID3B	Q8IVW6	1	2
VBP1	P61758	1	1
CFL1	P23528	1	
RNF113A	O15541	1	
DMD	P11532		1
IGHG1	P01857		1
SPRR2E	P22531		1
RBBP4	Q09028	2	
HMGB3	O15347	1	1
JUN	P05412	1	1
SUB1	P53999	1	1
DACH1	Q9UI36	1	1
RUVBL1	Q9Y265	1	1
GSDMA	Q96QA5	1	1
HNRNPA2B1	P22626	1	
FLYWCH2	Q96CP2	1	1
ALDOA	P04075	2	
MDC1	Q14676	1	
C19orf43	Q9BQ61	1	
CHMP2B	Q9UQN3	1	
PPM1G	O15355	1	
FAM50A	Q14320	1	
IREB2	P48200		1
FLG2	Q5D862		1
SSBP1	Q04837	1	
SPDL1	Q96EA4		1
HSPA1A	P08107	2	
HSPD1	P10809	1	
GTF2A1	P52655		1
BRD4	O60885		1
RFC4	P35249		1
DFFA	O00273	1	
SMNDC1	O75940	1	
SSB	P05455	1	
CPA1	P15085	1	
EIF4B	P23588	1	
PSME3	P61289	1	
YBX1	P67809	1	

ERH	P84090	1	
EWSR1	Q01844	1	
ALYREF	Q86V81	1	
CCDC43	Q96MW1	1	
PUS7	Q96PZ0	1	
ANP32E	Q9BTT0	1	
API5	Q9BZZ5	1	
TCERG1	O14776	1	
PEBP1	P30086	1	
ORC2	Q13416	1	
IMPDH2	P12268	1	
HNRNPA1	P09651	1	
ANP32A	P39687	1	
CDSN	Q15517		1
SUGP1	Q8IWZ8		1
CDV3	Q9UKY7	1	

APPENDIX 3: PUBLICATIONS

Adeno-associated virus Rep represses the human integration site promoter by two pathways that are similar to those required for the regulation of the viral p5 promoter.

Dutheil N, Smith SC, Agúndez L, Vincent-Mistiaen ZI, Escalante CR, Linden RM, Henckaerts E.

J Virology. 2014 Aug;88(15):8227-41. doi: 10.1128/JVI.00412-14. Epub 2014 May 14.

Manuscript Submitted:

Epigenetic regulation of adeno-associated virus latency.

Smith-Moore S, Neil S, Fraefel C, Linden RM, Rowe HM, Henckaerts E.

JVI Journals.ASM.org

Adeno-Associated Virus Rep Represses the Human Integration Site Promoter by Two Pathways That Are Similar to Those Required for the Regulation of the Viral p5 Promoter

Nathalie Dutheil,* Sarah C. Smith, Leticia Agúndez, Zoé I. Vincent-Mistiaen, Carlos R. Escalante,* R. Michael Linden, Els Henckaerts

Department of Infectious Diseases, King's College London School of Medicine, London, United Kingdom

ABSTRACT

Adeno-associated virus serotype 2 (AAV2) can efficiently replicate in cells that have been infected with helper viruses, such as adenovirus or herpesvirus. However, in the absence of helper virus infection, AAV2 establishes latency by integrating its genome site specifically into *PPP1R12C*, a gene located on chromosome 19. This integration target site falls into one of the most gene-dense regions of the human genome, thus inviting the question as to whether the virus has evolved mechanisms to control this complex transcriptional environment in order to facilitate integration, maintain an apparently innocuous latency, and/or establish conditions that are conducive to the rescue of the integrated viral genome. The viral replication (Rep) proteins control and direct every known aspect of the viral life cycle and have been shown to tightly control all AAV2 promoters. In addition, a number of heterologous promoters are repressed by the AAV2 Rep proteins. Here, we demonstrate that Rep proteins efficiently repress expression from the target site *PPP1R12C* promoter. We find evidence that this repression employs mechanisms similar to those described for Rep-mediated AAV2 p5 promoter regulation. Furthermore, we show that the repression of the cellular target site promoter is based on two distinct mechanisms, one relying on the presence of a functional Rep binding motif within the 5′ untranslated region (UTR) of *PPP1R12C*, whereas the second pathway requires only an intact nucleoside triphosphate (NTP) binding site within the Rep proteins, indicating the possible reliance of this pathway on interactions of the Rep proteins with cellular proteins that mediate or regulate cellular transcription.

IMPORTANCE

The observation that repression of transcription from the adeno-associated virus serotype 2 (AAV2) p5 and integration target site promoters is mediated by shared mechanisms highlights the possible coevolution of virus and host and could lead to the identification of host factors that the virus exploits to navigate its life cycle.

Adeno-associated virus serotype 2 (AAV2) is a human DNA virus that is dependent upon a number of factors provided by helper viruses in order to replicate efficiently (1–4). In the absence of helper factors, AAV2 has the ability to establish latency by site-specifically integrating its genome into chromosome 19 (5, 6). A complex interplay between host cellular, AAV2, and helper virus proteins leads to the tight regulation of a life cycle that is unique among eukaryotic viruses. The nonstructural proteins Rep78, Rep68, Rep52, and Rep40, encoded by the AAV2 *rep* genes, play a major role in orchestrating the different aspects of the AAV2 life cycle. The large Rep proteins, Rep78 and Rep68, transcribed from the p5 promoter, are multifunctional proteins with DNA binding, endonuclease, and helicase activities that control replication, integration, and transcription (reviewed in reference 7). The small Rep proteins, transcribed from the p19 promoter, share the nucleoside triphosphatase (NTPase) and helicase activity with the large Rep proteins and are required for efficient packaging of the viral particles (8). The Rep proteins have a distinctive autoregulatory role in that they control p5 and p19 transcription as well as p40-controlled transcription of the structural proteins. In the presence of helper virus, repression of the p5 promoter by Rep and cellular factors YY1 and MLTF is lifted (9–11), which leads to transactivation of transcription from all three promoters (12) but is regulated so that p40 transcript levels are higher than the p5 and p19 transcript levels (13). Sequences within, as well as outside, the viral promoter regions have been shown to be involved in Rep activation (14). Interestingly, during productive infection, the Rep proteins can mediate both activation and repression of transcription (15). More specifically, the large Rep proteins activate transcription from the p5 promoter through binding to the Rep binding site (RBS) present in the inverted terminal repeat (ITR) and mediate p5 repression by binding to the RBS in the p5 promoter. This repression can be partially lifted by the small Rep proteins and contributes to an autoregulatory loop, which maintains constant ratios of the p5 and p19 transcripts (16).

In the absence of helper virus, p5, p19, and p40 transcription is significantly reduced, leading to minute levels of Rep protein expression (17–19). Rep-mediated repression of p5 transcripts appears to be dependent on the NTP-binding motif present in the

Received 13 February 2014 **Accepted** 6 May 2014

Published ahead of print 14 May 2014

Editor: M. J. Imperiale

Address correspondence to Els Henckaerts, els.henckaerts@kcl.ac.uk.

* Present address: Nathalie Dutheil, Université de Bordeaux, Institut des Maladies Neurodégénératives and CNRS Institut des Maladies Neurodégénératives, UMR 5293, Bordeaux, France; Carlos R. Escalante, Department of Physiology and Biophysics, Virginia Commonwealth University School of Medicine, Richmond, Virginia, USA.

doi:10.1128/JVI.00412-14

central domain of the Rep proteins, as well as the presence of an intact RBS motif in the p5 promoter (20). In contrast, Rep-mediated repression of the p19 promoter requires only the NTP-binding motif of Rep (19). Transcriptional repression by Rep is not exclusive to AAV2 promoters but has also been observed for heterologous promoters, such as the HIV long terminal repeat (LTR), the human papillomavirus 18 (HPV18) upstream regulatory region (URR), and the major late transcription promoter of adenovirus (AdMLP), and is dependent on the Rep NTP-binding domain (21, 22). Rep also affects the expression of cellular genes, such as those encoding c-myc and c-sis/platelet-derived growth factor B (23–25); however, the significance of Rep-mediated regulation of these promoters in the context of AAV2's life cycle has yet to be established.

AAV2 has the ability to site-specifically integrate its genome into a gene, *PPP1R12C*, that encodes a protein that is thought to be a component of the regulatory subunit of myosin light chain phosphatase (26). Analysis of the 5′ untranslated region (UTR) of this gene led to the observation that the minimal promoter region of *PPP1R12C* and the p5 promoter have two important *cis*-regulatory elements in common, namely, the RBS and YY1 sites (27). This observation, together with the fact that AAV2 establishes latency by integrating into a ubiquitously transcribed locus, provoked the question of whether AAV2 has evolved the ability to control transcription of the AAV2 integration target locus in order to aid the establishment of latency and secure viral rescue.

We demonstrate here that Rep proteins efficiently repress expression from the target site *PPP1R12C* promoter, employing mechanisms similar to those described for Rep-mediated viral p5 promoter regulation. We provide evidence that the observed repression is based on two distinct mechanisms: one relies on the presence of a functional Rep binding motif within the 5′ UTR of *PPP1R12C*, whereas the second pathway requires only an intact NTP-binding site within the Rep proteins, possibly indicating the reliance of this pathway on interactions of the Rep proteins with factors of or associated with the cellular transcription machinery. We propose that the p5 promoter has coevolved with the host *PPP1R12C* promoter, thereby ensuring the possibility that Rep can control the expression at the viral integration site.

MATERIALS AND METHODS

Cell lines and viruses. HEK-293T/17 human embryonic kidney cells (ATCC CRL-11268) and HeLa human cervical epithelial cells (ATCC CCL-2) were cultured in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen) supplemented with 10% fetal bovine serum (FBS) (Invitrogen).

AAV2 and human adenovirus type 5 (Ad5) were produced and purified as previously described (28).

Infections. Cells were infected in 60-mm plates at 70% confluence by adding increasing amounts of wild-type (wt) AAV2 (multiplicity of infection [MOI], 1 to 10^4 infectious units per cell) to 1 ml of DMEM and 10% FBS. After 2 h of incubation at 37°C, 293T and HeLa cells were coinfected with adenovirus type 5 at an MOI of 5 and 10 PFU, respectively. After 1 h of incubation at 37°C, the inoculum was removed, and 3 ml of fresh medium was added to the cells.

Reporter constructs. To increase the sensitivity of the reporter construct, we replaced DsRed2.1 with mCherry in the previously used −332/+94 plasmid (pND26). The plasmid contains the *PPP1R12C* promoter region from nucleotides (nt) −332 to +94 relative to the transcription start site (27).

The PPP1R12Cp-mCherry (pND203) reporter construct was cloned in two steps. The mCherry open reading frame (ORF) was first amplified by PCR using ND381 (5′ GGATCCCACAACCATGGTGAGCAAGGGC-3′) and ND382 (5′-GCGGCCGCTACTTGTACAGCT-3′) as primers and pTW149 as the template (28). The resulting PCR product was cloned into pCR2.1 (Invitrogen) to create pND202. To generate the PPP1R12Cp-mCherry (pND203) and p5-mCherry (pND208) reporter constructs, the mCherry sequence linked to a BamHI/NotI fragment from pND202 was subcloned into the BamHI/NotI-digested pND26 and p5 vectors (pND85), respectively (27). The p5-mCherry construct contains the p5 promoter region (nt 190 to nt 310 relative to the AAV2 sequence) described by Cheung et al. (29).

Mutations within the *PPP1R12C* RBS were introduced by inserting annealed primers into the BamHI/SmaI sites of the pND23 plasmid (the *PPP1R12C* promoter region from nt −332 to +20 cloned into pBluescript II SK [Fermentas]) (27). The resulting plasmid, pND31 (the *PPP1R12C* promoter region from nt −332 to +94), was digested with BamHI and HindIII, and the *PPP1R12C* promoter fragment containing the mutated RBS motif was blunt ended prior to ligation into the SmaI site of the pDsRed2.1 vector (Clontech) to create plasmid pND34 (PPP1R12Cp-RBS*-DsRed2.1). To generate the PPP1R12Cp-RBS*-mCherry reporter construct (pND212), the mCherry-containing BamHI/NotI fragment from pND202 was subcloned into the BamHI/NotI sites of the pND34 plasmid. The same mutation was introduced into the p5 RBS by site-directed mutagenesis on the pND84 (p5-pCR2.1) template (27) to create pND235. To generate the p5-RBS*-mCherry construct (pND236), the SmaI fragment containing the p5 promoter region from pND235 was inserted into the SmaI sites of pND208.

To generate the pEGFP-PPP1R12Cp-mcherry bidirectional promoter construct (pND238), the blunt-ended AflII/MscI enhanced green fluorescent protein (EGFP)-encoding fragment of pIRES2-EGFP (Clontech) was subcloned into the blunt-ended EcoRI site of pND203.

Rep-expressing constructs. All Rep-expressing constructs were cloned into the pIRES2-EGFP vector. To generate the Rep78-expressing plasmid (pR78-IRES2-EGFP), the Rep78-encoding DraI/SphI fragment from plasmid pHis-Rep78 (30) was first ligated to annealed adaptors and inserted between the SmaI and NheI sites of the pIRES2-EGFP vector. To generate pR78Y156F-IRES2-EGFP, the SacII/SalI fragment containing the Y156F mutation from plasmid pHis-Rep68Y156F (30) was cloned into the SacII/SalI-digested pR78-IRES2-EGFP vector. To generate pRep68Y156F-IRES2-EGFP (pND21), the BamHI fragment of pHis-Rep68 (31) containing the C-terminal region of Rep68 was inserted between the BamHI sites of pR78Y156F-IRES2-EGFP. The initiation methionine of Rep52 and Rep40 was mutated to a glycine (M225G) by site-directed mutagenesis. The resulting PCR fragment was cloned into the pCR2.1 vector to generate pND102. The SacI/SalI fragment containing the M225G mutation from pND102 was inserted into the corresponding SacI/SalI fragment of pRep78Y156F-IRES2-EGFP and pND21 to generate pR78Y156F/M225G-IRES2-EGFP (pND105) and pR68Y156F/M225G-IRES2-EGFP (pND104), respectively.

The 5′ ends of Rep52 and Rep40 were generated by PCR on a pAV2 template with primers ND144 (5′-GATATCGCACAACATGGAGCTGG TCGGG-3′) and ND67 (5′-CATCCGGTCTTGCAACGGCTGC-3′), and the resulting PCR product was subcloned into the pCR2.1 cloning vector to create pND20. pRep52-IRES2-EGFP (pND22) was constructed by inserting the 5′ end of Rep52 containing the EcoRV/SalI fragment from pND20 into the EcoRV/SalI sites of pRep78Y156F-IRES2-EGFP. To generate pRep40-IRES2-EGFP (pND56), the Rep40 3′-end BamHI fragment from pND21 was inserted into the BamHI sites of pND22.

One, two, or three copies of the simian virus 40 (SV40) large T antigen nuclear localization signal (NLS), PKKKRKV, were added in frame to the 3′ end of the N208 open reading frame. PCR fragments amplified on pRep78Y156F-IRES2-EGFP with primers containing the NLS sequences were subcloned into pCR2.1 to generate pND59 (pN208Y156F-NLS1) and pND75 (pN208Y156F-NLS2). pND81 (pN208Y156F-NLS3) was generated by PCR on a pND75 template. To generate pN208Y156F-NLS_1-IRES2-EGFP (pND62), pN208Y156F-NLS_2-IRES2-EGFP (pND82#2),

and pN208Y156F-NLS_3-IRES2-EGFP (pND82#3), the PstI/BamHI fragment containing the 3′ end of N208 from pND59, pND75, and pND81 was inserted into the corresponding PstI/BamHI sites of pRep78Y156F-IRES2-EGFP.

To generate pRep68Y156F/K340H-IRES2-EGFP (pND25), the SacI/AccI fragment of pHis-Rep68K340H (32, 33), containing the K340H mutation, was inserted between the SacI/AccI sites of pND21. pRep68Y156F/M225G/K340H-IRES2-EGFP (pND140) and pRep40/K116H-IRES2-EGFP (pND80) were generated by inserting the BamHI fragment from pND25, containing the K340H mutation, into the corresponding BamHI fragments from pND104 and pND56. pRep78Y156F/M225G/K340H-IRES2-EGFP (pND146) and pRep52K116H-IRES2-EGFP (pND143) were generated by inserting a fragment containing the K340H mutation from pND140 (SacI/AccI) and pND80 (EcoRV/SalI) into the corresponding sites of pRep78Y156F-IRES2-EGFP.

To delete the internal ribosome entry site 2 (IRES2)-EGFP sequence from Rep-expressing vectors and generate pΔIRES2-EGFP (pND216), the plasmid pIRES2-EGFP was digested with BamHI and NotI, and the overhangs were blunt ended with the Klenow DNA polymerase prior to religation. The Rep78 and Rep68 ORFs were amplified by PCR from the pND105 and pND104 plasmids using the SmaI-containing primers ND387 (5′-CCCGGGATATC GCACAACATGCCGGGG-3′) and ND388 (5′-CCCGGGTTATTGTTCAA AGATGCAGTCATCCAAATC-3′), and ND387 and ND389 (5-CCCGGGT-CAGAGAGAGTGTCCTCGAGC-3′), respectively. The resulting PCR fragments were subcloned into the PCR2.1 cloning vector to generate pND221 (Rep78Y156F/M225G-pCR2.1) and pND220 (Rep68Y156F/M225G-pCR2.1). pRep78Y156F/M225G-ΔEGFP (pND227) and pRep 68Y156F/M225G-ΔEGFP (pND226) were generated by subcloning the Rep-encoding SmaI fragment from pND221 and pND220 into the SmaI site of pND216. pRep52-ΔEGFP (pND230) and pRep40-ΔEGFP (pND229) were generated by inserting the 5′ ends of the Rep52 and Rep40 EcoRV/BstEII fragments from pND22 and pND56 into the EcoRV/BstEII sites from pND227 and pND226, respectively. Plasmids pRep78Y156F/M225G/K340H-ΔEGFP (pND231), pRep68Y156F/M225G/K340H-ΔEGFP (pND232), pRep52K116H-ΔEGFP (pND233), and pRep40K116H-ΔEGFP (pND234) were generated by inserting the EcoRV/BstEII fragments containing the K340H mutation from plasmids pND146, pND140, pND143, and pND80 into the corresponding EcoRV/BstEII sites of plasmids pND227, pND226, pND230, and pND229. All the vectors were sequenced.

Transient transfections. Transient transfections were performed in 60-mm plates using Lipofectamine 2000 according to the manufacturer's instructions (Invitrogen). At 70% confluence, 293T cells were transfected with 4.5 μg of reporter plasmid and 9 μg of Rep-expressing plasmid. Cells were harvested 48 h posttransfection and assayed for plasmid DNA uptake and RNA and protein levels.

Plasmid uptake determination. The transfected cells were lysed in 0.2 M NaOH and 10 mM EDTA, boiled for 15 min at 90°C, and loaded onto a Hybond nitrocellulose membrane (Amersham) using a slot blot manifold (Bio-Rad). The membranes were hybridized to mCherry or EGFP probes to estimate the amount of reporter plasmid taken up by the cells. The probes were generated by PCR using 5′-GGATCCCACAACCATGG TGAGCAAGGGC-3′ and 5′-GCGGCCGCTACTTGTACAGCT-3′ for mCherry and 5′-GCTAGCCACAACCATGGTGAGCAAGGGC-3′ and 5′-GCTAGCTTACTTGTACAGCTCGTCCATGCCG-3′ for EGFP. Transfection efficiencies were normalized to plasmid reporter uptake.

Real-time qRT-PCR. Total RNA was extracted with the RNeasy kit (Qiagen) and reverse transcribed using SuperScriptIII reverse transcriptase (RT) (Life Technologies). Real-time quantitative RT-PCR (qRT-PCR) was performed on 50 to 100 ng cDNA using TaqMan Gene Expression Assays for *ACTB* (Hs99999903_m1) and *PPP1R12C* (Hs01085952_m1) and TaqMan Universal PCR master mix (Life Technologies). Relative expression levels were determined by the comparative threshold cycle (C_t) method (34).

Northern blot analysis. Total RNA was extracted with the RNeasy kit (Qiagen). Ten micrograms (infected cells) or 2 μg (transiently transfected cells) of RNA was separated on a 1.2% formaldehyde-agarose gel and transferred onto a nitrocellulose membrane. The membranes were first hybridized to [α-^{32}P]dCTP-labeled mCherry, PPP1R12C (exons 18 to 22), rep, or GFP probes; stripped at 65°C in 50% formamide, 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate); and rehybridized to a β-actin cDNA probe. The PPP1R12C and β-actin cDNA probes were generated as described by Dutheil et al. (27). The mCherry and GFP cDNA probes were generated by PCR using the primers mentioned above. The rep probe was generated by PCR on a plasmid containing a 315-bp PstI/SacI fragment from pAV2 using 5′-GGATCCTCAATTCTGATTCTCTT TG-3′ and 5′-CCCGGGGGTCCTGTATTAGAGGTCACGTG-3′. All Northern blots were analyzed with a Typhoon PhosphorImager (Molecular Dynamics) and then exposed to an X-ray film to generate high-quality images. ImageQuant TL software was used to calculate fold repression. Each average repression level is represented as the mean and standard error of the mean (SEM).

Western blot analysis. Forty-eight hours after transfection, cells were lysed in RIPA buffer (50 mM Tris-HCl [pH 8], 150 mM NaCl, 0.1% SDS, 1% Nonidet P-40, 0.5% sodium deoxycholate, 1× Complete protease inhibitor cocktail [Roche Applied Science]), and proteins were quantified by the bicinchoninic acid (BCA) protein assay (Pierce). For each condition, the same amount of protein (10 μg) was separated on a 15% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane (Hybond-C Extra nitrocellulose; Amersham Biosciences). The membranes were blocked with 5% nonfat dry milk in Tris-buffered saline (TBS) (100 mM Tris-Cl, pH 7.5, 150 mM NaCl) containing 0.5% Tween 20 (TBS-T) for 1 h at room temperature and incubated with primary antibody overnight at 4°C. The membranes were washed in TBS-T buffer (three 15-min washes) and then incubated with secondary antibody conjugated to horseradish peroxidase for 1 h at room temperature. After three washes in TBS-T buffer, the membranes were developed using enhanced-chemiluminescence (ECL) substrate (Pico detection kit; Pierce). Using an ImageQuant LAS 4000 Biomolecular Imager and ImageQuant software (GE Healthcare Life Sciences), band densitometry was performed, and the result was normalized against the value of actin protein expression. After visualization of the desired protein, the membranes were stripped in Restore Western blot stripping buffer (Pierce) for 30 min at room temperature. The membranes were washed four times in TBS-T buffer, blocked with 5% nonfat dry milk, and then hybridized with specific antibody.

The primary antibodies used in the study were antibodies against mCherry (1:16,000 dilution in 1% bovine serum albumin [BSA] in TBS-T; Clontech; rabbit polyclonal red fluorescent protein [RFP] antibody; catalog no. 632397); GFP (1:5,000 dilution in 5% milk in TBS-T; Roche Applied Science; mouse monoclonal GFP antibody; catalog no. 11 814 460 001); AAV2 Rep proteins Rep78, Rep68, Rep52, and Rep40 (1:100 dilution in 5% nonfat dry milk–PBS-T for monoclonal antibody clone 303.9 [Progen Biotechnik catalog no. 61069] or 1:500 dilution in 1% BSA–PBS-T for monoclonal antibody clone 226-7 [Acris Antibodies catalog no. BM5012SU]); AAV-2 Rep proteins Rep78, Rep68, and Rep78-ΔN208 (1:200,000 dilution in 5% nonfat dry milk–PBS-T; rabbit polyclonal anti-N208 antibody); and actin proteins (1:10,000 dilution in 5% nonfat dry milk in PBS-T; BD Biosciences; mouse monoclonal actin antibody; catalog no. 612656). Polyclonal anti-N208 antibody was produced in rabbits immunized with the truncated Rep protein Rep78-ΔN208, containing the first N-terminal 208 amino acids of Rep78 and Rep68 (Cocalico Biologicals Inc.).

Goat anti-mouse (Jackson ImmunoResearch Laboratories; catalog no. 115-035-003) or anti-rabbit (Jackson ImmunoResearch Laboratories; catalog no. 111-036-003) secondary antibody conjugated to horseradish peroxidase was used at a dilution of 1:10,000 in 5% nonfat dry milk in PBS-T.

Each average repression level is represented as the mean and SEM.

Immunofluorescence. HeLa cells were grown in 24-well plates on coverslips coated with poly-L-lysine. At 50% confluence, the HeLa cells were transfected with 0.75 μg of DNA using the Lipofectamine Plus re-

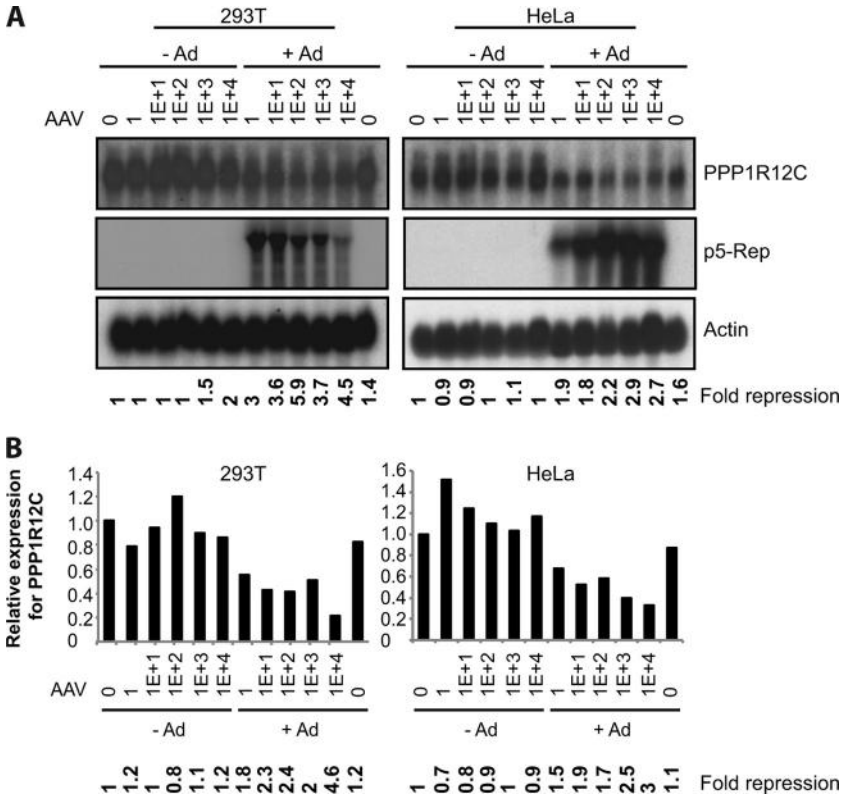


FIG 1 *PPP1R12C* and *rep* expression levels in AAV2-infected cells. (A) Representative example of Northern blot analysis of 293T (left) and HeLa (right) cells infected with wt AAV2 at increasing MOIs (1 to 10,000) in the absence (−Ad) and presence (+Ad) of adenovirus 48 h postinfection. Northern blots were hybridized with PPP1R12C (top), p5-Rep (middle), and β-actin (bottom) cDNA probes. Relative fold repression of *PPP1R12C* expression is indicated at the bottom. (B) Determination of *PPP1R12C* expression levels by real-time qRT-PCR in the samples shown in panel A confirms downregulation of *PPP1R12C* expression during productive AAV2 infection in two different cell lines.

agents (Invitrogen). The cells were fixed 48 h posttransfection in cold acetone, permeabilized with 0.1% Triton X-100 for 5 min, and blocked in 5% normal goat serum overnight. The slides were incubated with primary rabbit anti-N208 antibody (1:1,000 dilution in PBS) for 1 h at room temperature, washed in PBS, and incubated with a Cy3 goat anti-rabbit antibody (1:1,000 dilution in PBS; Jackson ImmunoResearch; catalog no. 711-165-152) for 1 h at room temperature. The slides were subsequently washed in PBS and mounted in Vectashield Mounting Medium containing DAPI (4′,6-diamidino-2-phenylindole; Vector Laboratories). Images were acquired at ×100 magnification using a Leica DMRA2 fluorescence microscope with a Hamamatsu charge-coupled-device (CCD) digital camera and analyzed with Openlab software (Improvision).

Fluorescence-activated cell sorting. To determine the effect of Rep on endogenous *PPP1R12C* expression, 293T cells transfected with the different Rep-IRES-GFP expression constructs were harvested 48 h posttransfection and sorted for GFP expression using the BD FACSAria II (Becton Dickinson) prior to RNA isolation.

RESULTS

Under permissive conditions, AAV2 infection leads to downregulation of *PPP1R12C* expression. To determine if AAV2 infection leads to changes in the expression levels of the target site *PPP1R12C* promoter, 293T and HeLa cells were infected with wt AAV2 using increasing MOIs in the presence and absence of adenovirus (wt Ad5). Coinfection with adenovirus leads to permissive conditions, which support efficient AAV2 replication, whereas infection with AAV2 alone represents nonpermissive conditions under which the virus can establish latency. Total RNA was extracted 48 h postinfection (p.i.), and the expression of *PPP1R12C* was analyzed by Northern blotting and real-time qRT-PCR. In the absence of adenovirus infection, expression levels of *PPP1R12C* were similar in control and wt AAV2-infected HeLa cells at all MOIs tested (Fig. 1A, top right). In 293T cells, the same tendency was observed; however, at higher MOIs, slightly lower levels of *PPP1R12C* expression were observed than in control cells (Fig. 1A, top left). In contrast to nonpermissive conditions, *PPP1R12C* expression is clearly repressed in wt AAV2- and adenovirus-coinfected cells compared to control cells, and this was observed for both cell lines (Fig. 1A, top row). Importantly, this effect does not appear to be mediated by adenovirus alone, as *PPP1R12C* expression levels were similar in control and adenovirus-infected cells. Note that the actin expression levels remained unaltered upon productive AAV2 infection. The observed repression of *PPP1R12C* expression under permissive conditions was confirmed by real-time qRT-PCR (Fig. 1B).

As the large Rep proteins, Rep78 and Rep68, are highly expressed under permissive conditions, we determined whether downregulation of *PPP1R12C* expression might be related to the presence of Rep78 or Rep68 (Rep78/Rep68) transcripts (12). As expected, the AAV2 p5 transcripts can be detected in AAV2- and adenovirus-coinfected cells, whereas Rep78/Rep68 expression cannot be detected in AAV2- or adenovirus-infected or control cells (Fig. 1A, middle row). These data indicate that a decrease in *PPP1R12C* expression might be correlated with an increase in p5

transcript levels, suggesting that the AAV2 Rep proteins can modulate *PPP1R12C* expression within its genomic locus. It is interesting that in 293T cells, despite the decrease of Rep transcript levels at higher MOIs, *PPP1R12C* expression levels remain strongly repressed.

A potential link between downregulation of expression from the AAV2 integration target site promoter and Rep expression was further supported by a time course experiment of AAV2 and adenovirus infection in 293T cells, for which *PPP1R12C* and Rep expression levels were determined by real-time qRT-PCR. The presence of Rep proteins in AAV2- and adenovirus-coinfected cells was confirmed by Western blotting. Twenty-four hours after infection, we could observe a modest decrease in relative *PPP1R12C* expression levels in cells coinfected at a high MOI (Fig. 2A, left), which became more pronounced at later time points (Fig. 2A and B, left). At 48 h postinfection, we could also observe downregulation of expression in the absence of adenovirus. Interestingly, Rep transcripts were detected under all conditions but increased strongly when adenovirus was added to the cultures. In general, Rep transcript levels decreased as the infection progressed. Western blot analysis of the samples showed strong expression of the large Rep proteins 24 h after coinfection, increasing expression of the small Rep proteins at 30 h postinfection, and decreasing Rep protein levels at 48 h postinfection (Fig. 2A, B, and C, right).

Both large and small Rep proteins repress *PPP1R12C* promoter activity. In order to test whether the observed repression is directly mediated by the viral Rep proteins, we transfected 293T cells with different Rep expression constructs. Since the chromosomal RBS and terminal resolution site (TRS) are located in the 5′ UTR of the *PPP1R12C* gene (27), it is conceivable that Rep78 and Rep68 can interfere with the *PPP1R12C* promoter activity by introducing a site-specific nick in the 5′ UTR (35, 36). Therefore, all experiments were executed using the Rep mutant RepY156F, which lacks nicking enzymatic activity (37) (Fig. 3A). Since the Rep52 and Rep40 proteins are the N-terminally truncated forms of Rep78 and Rep68, respectively, the initial methionine of Rep52 and Rep40 was mutated to a glycine to accomplish expression of the large Rep proteins only (20, 21) (Fig. 3A). We also examined whether the truncated protein containing only the first N-terminal 208 amino acids of Rep, Rep78-Y156F-ΔN208, could affect *PPP1R12C* promoter activity. The rationale for using the Rep78-Y156F-ΔN208 protein is that this truncated protein has previously been shown to be the minimal domain that can efficiently bind to both AAV2 and *PPP1R12C* RBS sequences *in vitro* while retaining its TRS endonuclease activity (30). Because Rep78-Y156F-ΔN208 lacks the bipartite NLS, which is localized in the C-terminal domain of Rep78/Rep68 (38), the C-terminal region of Rep78-Y156F-ΔN208 was fused to either one, two, or three tandem repeats of the NLS of SV40 large T antigen. In order to assess whether Rep78-Y156F-ΔN208 is effectively translocated to the nucleus, we performed immunofluorescence microscopy on cells transfected with the different NLS constructs. As shown in Fig. 3B, Rep78-Y156F-ΔN208 containing three copies of the SV40 large T antigen NLS is, like wild-type Rep78, mainly located in the nucleus. Rep78-Y156F-ΔN208 containing only one or two copies of the SV40 large T antigen NLS is predominantly located in the cytoplasm and in the perinuclear region (data not shown). All Rep variants were cloned into an IRES-GFP vector. The transfected cells were sorted for GFP expression to ensure the presence of Rep. As shown in Fig. 3C, all Rep proteins except Rep78-Y156F-ΔN208 had the ability to strongly repress endogenous *PPP1R12C* expression levels in GFP-positive cells (Fig. 3C). Western blot analysis confirmed expression of the different Rep proteins in the transfected cells (Fig. 3D).

In order to obtain better insight into potential mechanisms responsible for the observed Rep-mediated repression of *PPP1R12C* expression, we cotransfected Rep constructs (Fig. 3A) with reporter constructs containing the mCherry gene under the control of the *PPP1R12C* or AAV2 p5 promoter (Fig. 4A). Since Rep proteins can modulate the expression of a number of cellular and viral genes, including the cytomegalovirus (CMV) promoter (23, 24, 39), it is not possible to correct for variations in transfection efficiency by normalization to the activity of a cotransfected plasmid expressing a marker gene under the control of a viral or eukaryotic promoter. Therefore, transfection efficiencies were normalized to plasmid DNA uptake, as described previously (39). Figure 4B shows a representative example of the Northern and Western blot analyses performed. Transcript and protein levels from three independent experiments are shown in Fig. 4C. Rep78 and Rep68 strongly inhibit *PPP1R12C* promoter-driven expression at both the RNA and protein levels. In contrast to the large Rep proteins, Rep52 and Rep40 reproducibly inhibit *PPP1R12C* promoter activity, but at lower levels. Changes in RNA levels parallel those observed for protein levels. These data demonstrate that all Rep proteins have the ability to mediate repression of the *PPP1R12C* promoter despite significant differences in the levels of inhibition between Rep78/Rep68 and Rep52/Rep40. The difference in ability to repress *PPP1R12C* expression between the small and large Rep proteins was not as evident with the endogenous gene, which can be explained by a lower gene copy number than in the overexpression system. Similarly to what is shown in Fig. 3C, the level of *PPP1R12C* expression in the presence of Rep78-Y156F-ΔN208 is comparable to that detected in the absence of Rep (Fig. 4B and C). These data indicate that Rep78-Y156F-ΔN208 does not have the ability to regulate *PPP1R12C* promoter activity, suggesting that the DNA binding domain is not sufficient to inhibit *PPP1R12C* expression.

As the AAV2 p5 promoter shares common regulatory elements with the *PPP1R12C* promoter (27), the question arises as to whether the viral Rep proteins might regulate the two promoter activities by similar mechanisms. Therefore, we compared Rep's effects on the p5 and *PPP1R12C* promoters. Interestingly, all Rep proteins downregulate the AAV2 p5 promoter similarly to what we observed for the *PPP1R12C* promoter (Fig. 4B and C, right). In agreement with previously published data (19), the large Rep proteins completely repress p5 promoter activity, whereas the small Rep proteins inhibit the p5 promoter to a lesser extent. Of note, Rep78-Y156F-ΔN208 moderately represses p5 transcriptional activity, whereas the *PPP1R12C* promoter activity remains relatively unchanged in the presence of this truncated Rep protein.

The NTP-binding motif is required for Rep52- and Rep40-mediated repression of the p5 and *PPP1R12C* promoters. Previous studies established that a residue within the NTP-binding domain of Rep78 (K340) is critical for the negative regulation of the p19, HIV LTR, and HPV18 URR promoters (21), raising the hypothesis that the same residue might be required for *PPP1R12C* promoter repression. As Rep52 and Rep40 mainly consist of an NTP-binding domain with ATPase and helicase activities (Fig. 5A), we used the smaller Rep proteins to determine whether mu-

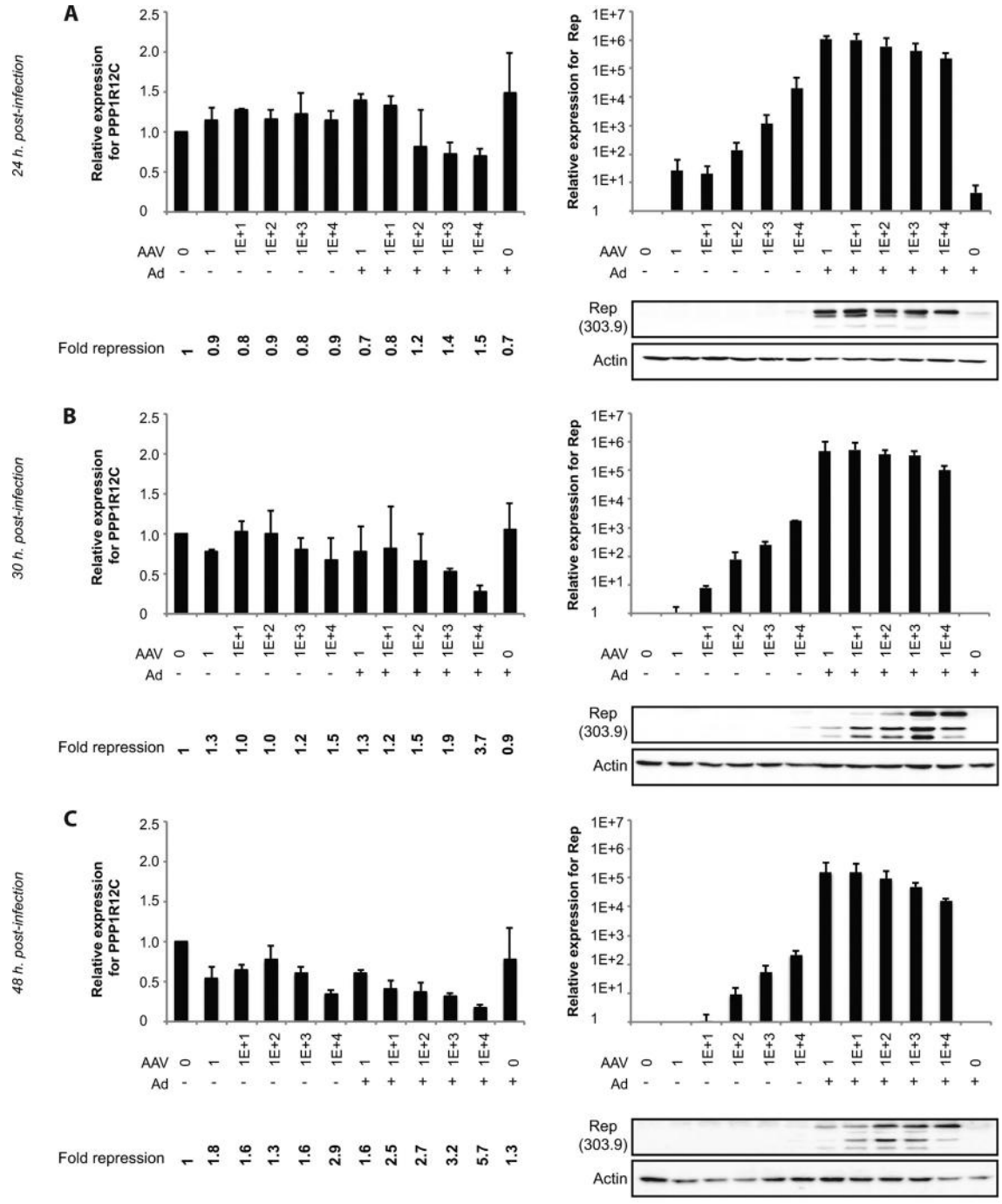


FIG 2 Time course of *PPP1R12C* and *rep* expression levels in AAV2-infected 293T cells. (A to C, left) Determination of *PPP1R12C* expression levels by real-time qRT-PCR in 293T cells 24 (A), 30 (B), and 48 (C) hours after AAV2 infection (at increasing MOIs in the absence and presence of adenovirus). Relative fold repression of *PPP1R12C* expression is indicated at the bottom. (A to C, top right) Determination of *rep* expression levels by real-time qRT-PCR in 293T cells 24 (A), 30 (B), and 48 (C) hours after AAV2 infection (at increasing MOIs in the absence and presence of adenovirus). Relative expression levels were determined in 2 independent infection experiments at each time point. (A to C, bottom right) Representative example of Western blot analysis of Rep expression in 293T cells 24 (A), 30 (B), and 48 (C) hours after AAV2 infection.

tations in the NTP-binding motif would have an effect on *PPP1R12C* and p5 promoter activities. To address this, wild-type or mutant Rep proteins were cotransfected with the *PPP1R12C* or p5 reporter construct. Northern and Western blot analyses were performed to determine RNA and protein expression levels, respectively (Fig. 5B). Transcript and protein levels from 3 independent experiments are shown in Fig. 5C. Similar to the data presented in Fig. 4B and C, wt Rep52 and wt Rep40 moderately inhibit the *PPP1R12C* and p5 promoter activity, as seen for RNA and protein levels (Fig. 5B and C). In contrast to the wt Rep pro-

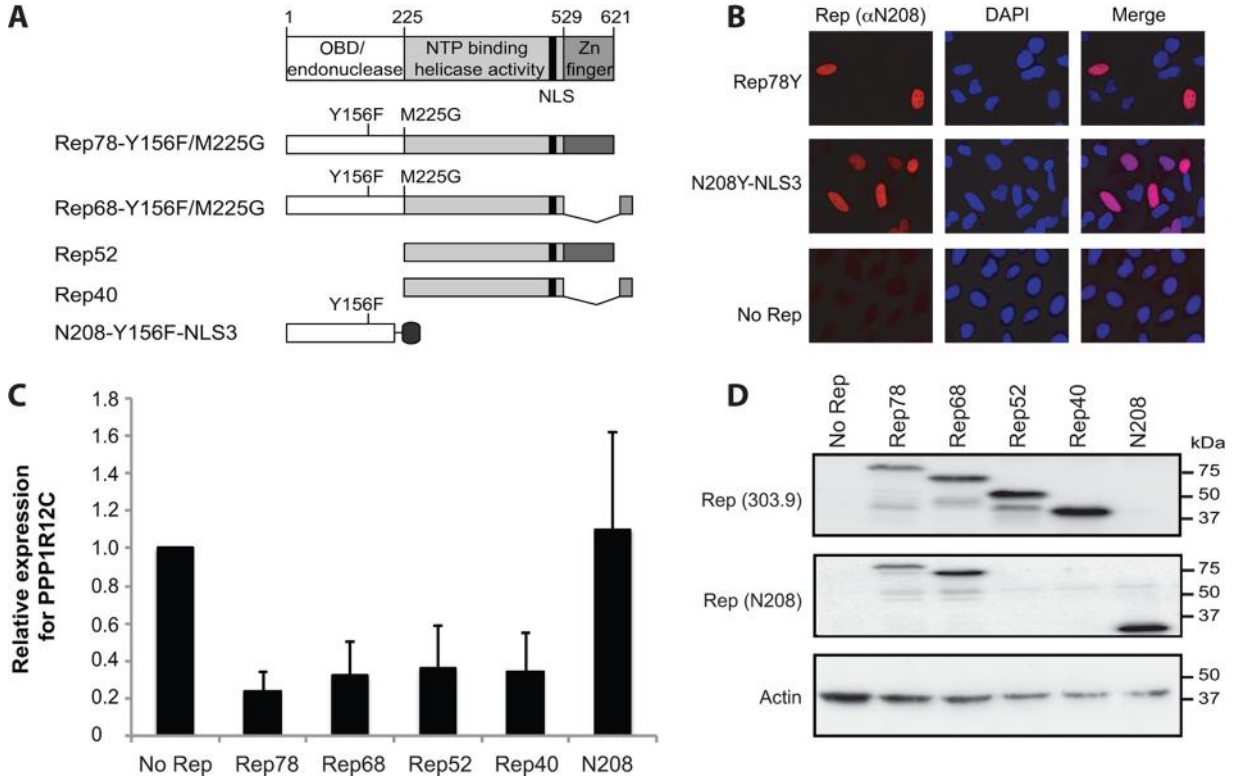
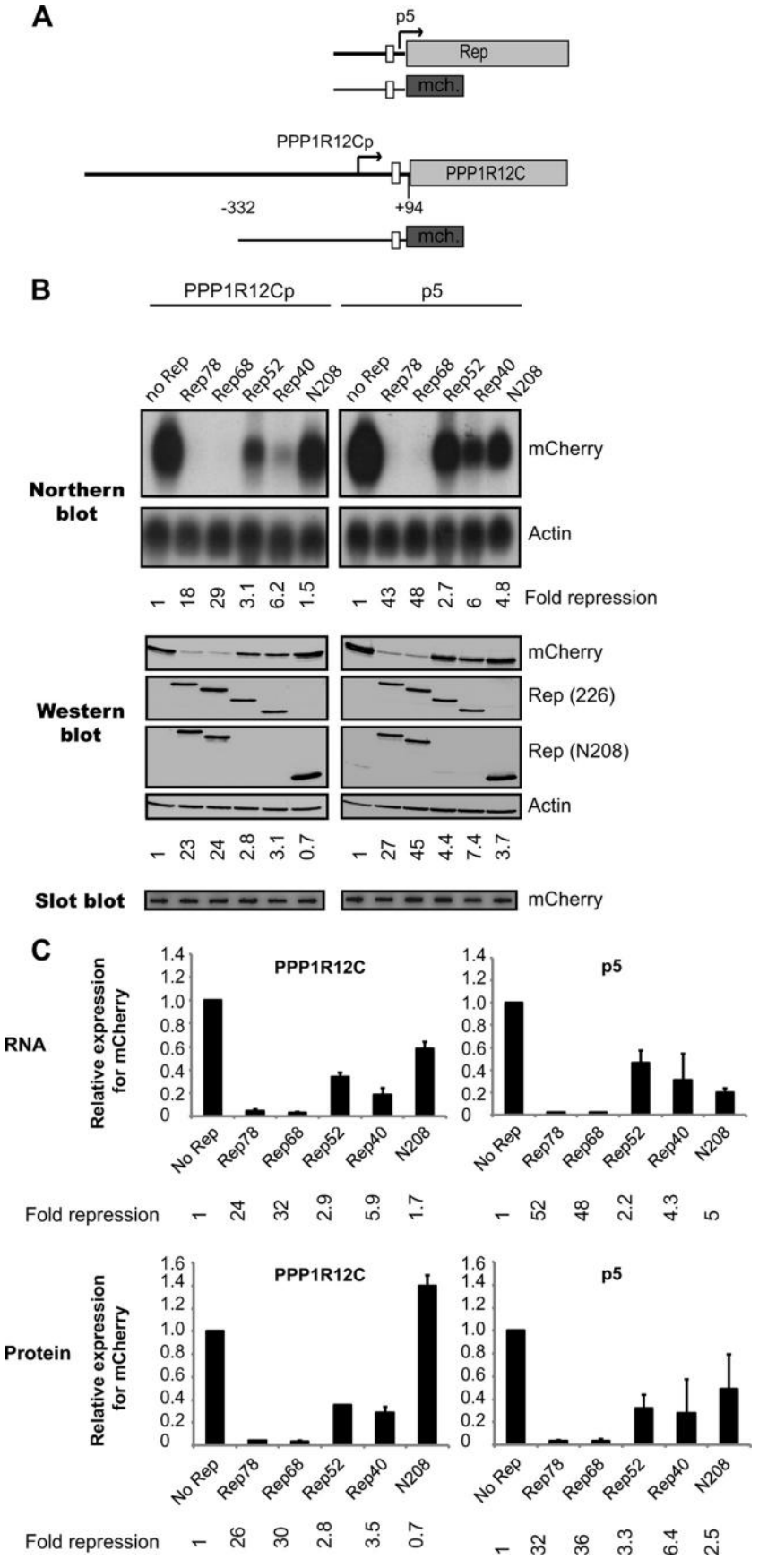


FIG 3 Endogenous *PPP1R12C* expression levels in Rep-transfected cells. (A) Schematic representations of the large (Rep78/Rep68), small (Rep52/Rep40), and truncated (N208) Rep proteins used for overexpression in 293T cells. At the top are shown the different functional domains present in Rep78. Nucleotide positions indicating the beginning and end of each domain are indicated above. The black bar shows the position of the NLS. Below are shown the large Rep proteins, Rep78 and Rep68, used for the analysis. Mutations were introduced to avoid endonuclease activity (Y156F) and simultaneous expression of the small Rep proteins (M225G). At the bottom are depicted the small Rep proteins (Rep52 and Rep40) and the N-terminally truncated Rep protein fused to 3 tandem repeats of the SV40 large T antigen NLS (N208-NLS3). The different Rep proteins were expressed from CMV promoter-IRES-GFP vectors. (B) Epifluorescence microscopy images of cells transfected with plasmids expressing Rep78-Y156F-N208Δ containing three copies of the SV40 large T antigen NLS and stained with DAPI and anti-Rep (α-Rep) (N208) antibody. The positive control consisted of cells transfected with a plasmid expressing full-length Rep78-Y156F; the negative control consisted of cells transfected with the CMV promoter-IRES-GFP vector. The images were taken at ×100 magnification. (C) Determination of *PPP1R12C* expression levels by real-time qRT-PCR in cells transfected with the different Rep constructs and sorted by fluorescence-activated cell sorter (FACS) for GFP expression. Relative expression levels were determined in 3 independent transfection experiments. The error bars indicate SEM. (D) Western blot analysis confirmed the presence of the respective Rep proteins in the transfected cells.

teins, the mutant Rep52-K116H and Rep40-K116H proteins, which harbor a mutation that corresponds to the K340H mutation in the large Rep proteins, have no clear effect on *PPP1R12C* and p5 promoter activity (Fig. 5B and C). The relief of repression by the NTP-binding mutants is not due to differences in protein expression levels, as they are similar for all Rep constructs, or even slightly higher for the Rep40-K116H mutant than for the corresponding wt protein, as indicated by Western blotting. In sum, these data demonstrate that Rep52 and Rep40 inhibit the *PPP1R12C* and p5 promoter activities in a similar manner and that this repression requires the consensus NTP-binding motif.

The RBS and the NTP-binding motif are both required for Rep78/Rep68-mediated repression of the p5 and *PPP1R12C* promoters. We next investigated the mechanism by which Rep78 and Rep68 mediate *PPP1R12C* repression. Since the AAV2 Rep78 and Rep68 proteins have the ability to bind to the RBS located within the 5′ UTR of the *PPP1R12C* gene (36), we hypothesized that, as observed for the repression of the p5 promoter (20), direct interaction of Rep with the RBS might also be involved in the inhibition of *PPP1R12C*. To address this, we used a reporter construct containing a mutation within the RBS that abolishes Rep binding to the RBS in the *PPP1R12C* promoter (36). We introduced the same mutation in the p5 RBS. Transcript and protein levels from 3 independent experiments were quantitated (Fig. 6C), and a representative example is shown in Fig. 6B. Similarly to what is shown in Fig. 4, Rep78- and Rep68-expressing plasmids strongly repress *PPP1R12C* and p5 promoter activities (Fig. 6B and C). Compared to the wt promoters, the mutation within the RBS motif reduces the ability of Rep78 and Rep68 to decrease the levels of *PPP1R12C* and p5 promoter transcripts (Fig. 6B and C). Although Rep's repressive effect on the mutant promoter is partially relieved, it is still clearly present. These data indicate that a direct interaction of Rep with the RBS is not sufficient to direct Rep-mediated repression of the *PPP1R12C* promoter, suggesting the existence of at least one additional mechanism at the basis of the observed phenomenon. It has previously been shown that Rep78 and Rep68 inhibit transcription from the p5 promoter by two different mechanisms. The first mechanism requires direct interaction of Rep with the RBS, while the second mechanism depends on the presence of a functional NTP-binding motif (20). We therefore tested the effects of the Rep78 and Rep68 nicking- and NTP-binding-negative mutant proteins (Rep78-Y156F-


A
p5
Rep
mch.
PPP1R12Cp
PPP1R12C
-332
+94
mch.
B
PPP1R12Cp
p5
no Rep
Rep78
Rep68
Rep52
Rep40
N208
Northern blot
mCherry
Actin
1 18 29 3.1 6.2 1.5
1 43 48 2.7 6 4.8
Fold repression
Western blot
mCherry
Rep (226)
Rep (N208)
Actin
1 23 24 2.8 3.1 0.7
1 27 45 4.4 7.4 3.7
Slot blot
mCherry
C
RNA
Relative expression for mCherry
PPP1R12C
p5
No Rep
Rep78
Rep68
Rep52
Rep40
N208
Fold repression
1 24 32 2.9 5.9 1.7
1 52 48 2.2 4.3 5
Protein
Relative expression for mCherry
PPP1R12C
p5
Fold repression
1 26 30 2.8 3.5 0.7
1 32 36 3.3 6.4 2.5

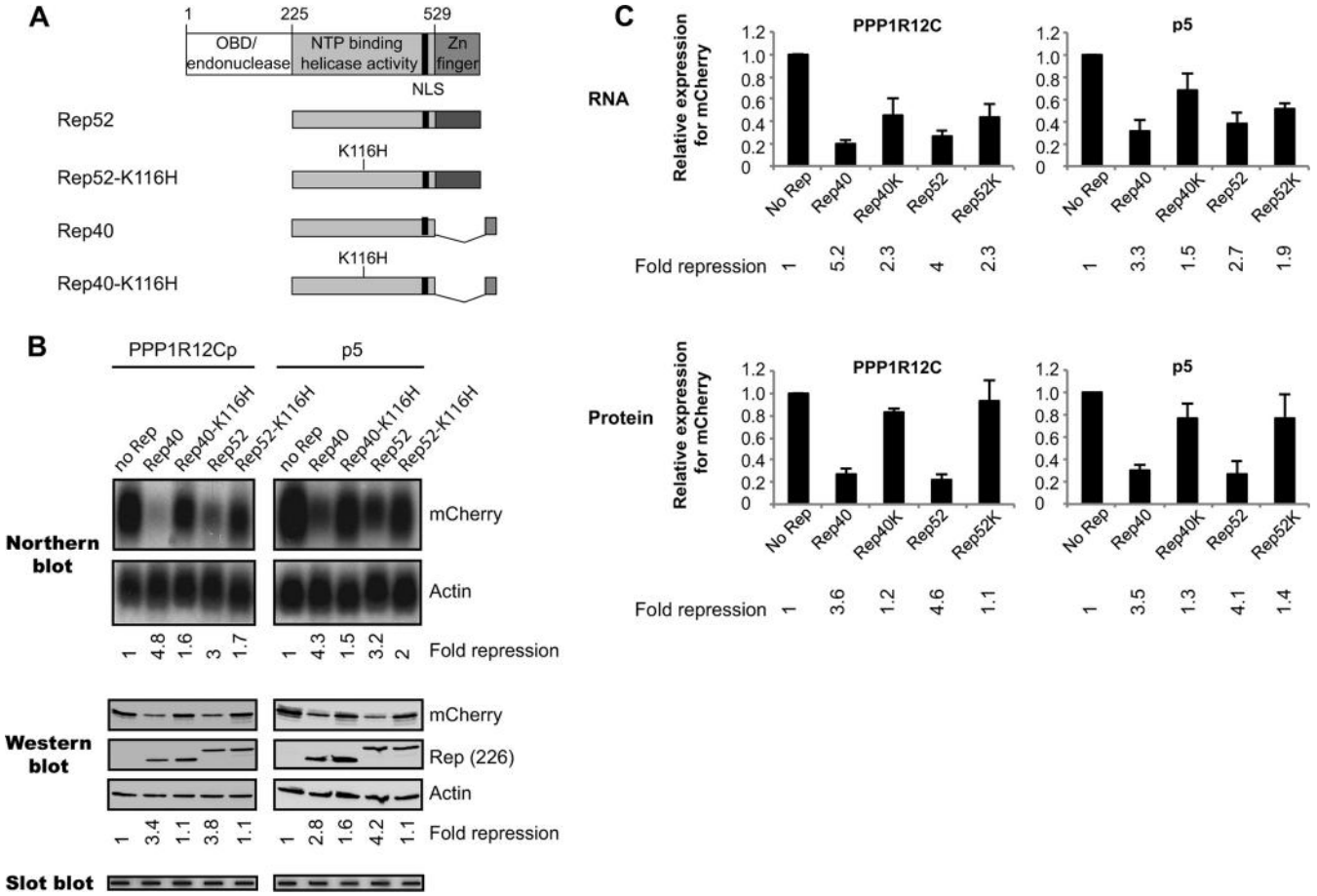


FIG 5 Analysis of the effects of the Rep52/Rep40 proteins and corresponding NTP-binding mutants on gene expression from the *PPP1R12C* and p5 promoters. (A) At the top are shown the different functional domains present in Rep78. Nucleotide positions indicating the beginning and end of each domain are indicated above. The black bar shows the position of the NLS. Below are shown schematic representations of the various small Rep proteins coexpressed with the *PPP1R12C* and p5 mCherry reporter plasmids. The NTP-binding mutants harbor a mutation at amino acid position 116 (K116H). (B) Northern blot analysis shows *PPP1R12C* (left) and p5 (right) transcription levels in 293T cells cotransfected with an mCherry reporter plasmid and various Rep-expressing plasmids. The blots were stripped and hybridized with a β-actin cDNA probe (lower blots). Relative fold repression of the mCherry transcription levels is indicated at the bottom of the blots. Western blot analysis shows the corresponding mCherry protein levels resulting from *PPP1R12C* (left) and p5 (right) promoter activity (top). Expression of the small Rep proteins and NTP-binding mutants was determined by hybridization with the 226-7 antibody. The blots were also incubated with a β-actin antibody, and relative fold repression was determined (indicated below the blots). Slot blot membranes hybridized to an mCherry probe show similar plasmid uptake for all experimental conditions. (C) Determination of average mCherry expression levels from the *PPP1R12C* (left) and p5 (right) promoters from 3 independent experiments by Northern and Western blot analyses. Average fold mCherry repression is indicated at the bottom of each graph. The error bars indicate SEM.

K340H and Rep68-Y156F-K340H) (Fig. 6A) on the wt and mutated *PPP1R12C* and p5 promoters. As shown in Fig. 6B and C (left), in the presence of an unaltered RBS, mutant Rep78/Rep68 proteins have a moderate repressive effect on the *PPP1R12C* promoter activity. This effect can also be observed for protein levels. Interestingly, the introduction of a mutation in the NTP-binding motif did not affect the repression of p5 transcription levels, as previously observed (19); however, a change in p5 repression was clear at the protein level (Fig. 6B and C, right). These results indicate that the activity of the NTP-binding motif in Rep78/Rep68 is

FIG 4 Analysis of the effects of the four AAV2 Rep proteins on *PPP1R12C* promoter (*PPP1R12C p*)- and AAV2 p5-directed gene expression. (A) Schematic representations of the AAV2 p5 and *PPP1R12C* mCherry reporter plasmids. The genomic structures of the Rep and *PPP1R12C* genes are depicted at the top. The bent arrows mark the p5 and *PPP1R12C* transcription start sites (TSS). The white boxes indicate the positions of the p5 RBS and *PPP1R12C* TRS-RBS motifs. (B) Northern blot analysis shows *PPP1R12C* (left) and p5 (right) transcription levels in 293T cells cotransfected with an mCherry reporter plasmid and the Rep-expressing plasmids shown in Fig. 3A. The blots were stripped and hybridized with a β-actin cDNA probe (lower blots). Relative fold repression of mCherry expression is indicated below the blots. Western blot analysis shows the corresponding mCherry protein levels resulting from *PPP1R12C* (left) and p5 (right) promoter activity. Rep expression was confirmed by Western blotting using the 226-7 and N208 antibodies. The blots were also incubated with a β-actin antibody, and relative fold repression was determined (indicated below the blots). Slot blot membranes hybridized to an mCherry probe show similar plasmid uptake for all experimental conditions. (C) Determination of average mCherry expression levels from the *PPP1R12C* (left) and p5 (right) promoters from 3 independent experiments by Northern and Western blot analyses. Average fold repression of mCherry expression is indicated at the bottom of each graph. The error bars indicate SEM.

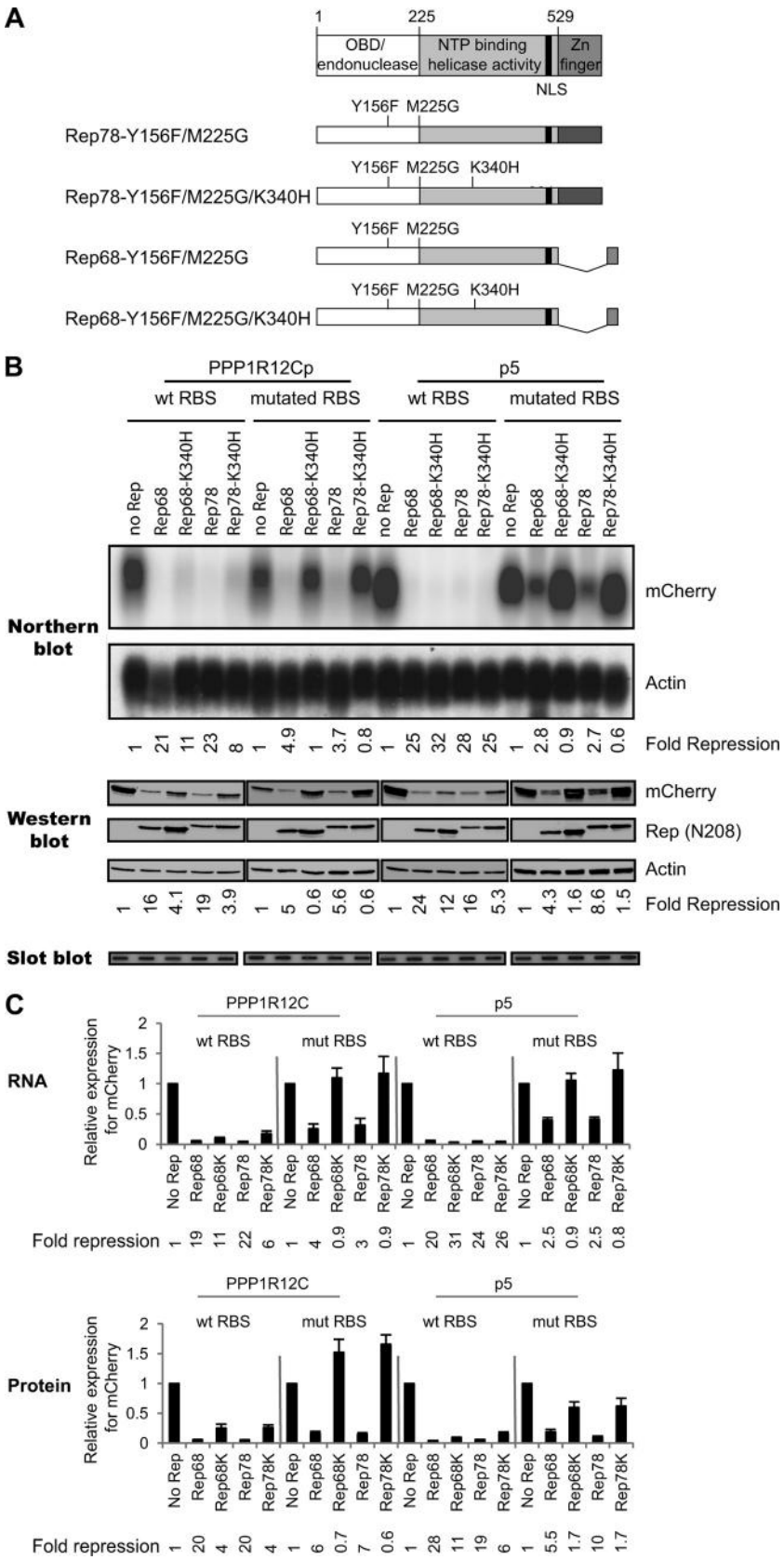


FIG 6 Analysis of the effects of the Rep78/Rep68 proteins and corresponding NTP-binding mutants on gene expression from wild-type and RBS mutant *PPP1R12C* and p5 promoters. (A) At the top are shown the different functional domains present in Rep78. Nucleotide positions indicating the beginning and end

not sufficient to accomplish complete *PPP1R12C* and p5 promoter repression. Therefore, we investigated whether the Rep78/Rep68 NTP-binding mutant proteins in the context of the mutated *PPP1R12C* and p5 promoters would lead to the complete abolishment of Rep78/Rep68's repressive effects. In contrast to the Rep78/Rep68-expressing plasmid, the NTP-binding-negative mutant was unable to repress transcription from the *PPP1R12C* and p5 promoters containing mutations in the RBS motif (Fig. 6B and C). Our data are in agreement with previously published results on Rep68-mediated repression of the p5 promoter (20). Taken together, our data provide strong evidence that Rep78/Rep68 exerts its negative regulatory effect on the *PPP1R12C* and p5 promoters by identical mechanisms. Rep-mediated repression of the p5 and *PPP1R12C* transcriptional activities requires, in addition to the NTP-binding motif, direct binding of Rep78/Rep68 to the RBS located within the promoter.

Rep-mediated repression of the antisense promoter requires only a functional NTP-binding motif. We have previously reported that the *PPP1R12C* promoter displays bidirectional promoter activities (27). The ability of Rep to regulate *PPP1R12C* transcription raises the question as to whether Rep is also able to regulate the transcriptional activity of the antisense *PPP1R12C* promoter. To address this question, the wt *PPP1R12C* promoter fragment used in the PPP1R12Cp-mCherry reporter construct, which exhibits both sense and antisense promoter activities (27), was cloned into a dual-reporter construct carrying the mCherry and EGFP genes in opposite directions. The *PPP1R12C* promoter drives the expression of mCherry, while the antisense promoter drives the expression of the EGFP cDNA. In order to simultaneously compare the effects of Rep on *PPP1R12C* sense and antisense promoter activities, the EGFP cDNA was deleted from all Rep-expressing constructs.

293T cells were transiently cotransfected with the *PPP1R12C* bidirectional reporter vector and various Rep-expressing constructs. In the first set of experiments, we investigated the effects of wt and NTP-binding mutant proteins on *PPP1R12C* promoter activity in the context of the bidirectional construct (Fig. 7A). Similarly to what is shown in Fig. 4, Rep78 and Rep68 strongly repress *PPP1R12C* transcription from the bidirectional promoter, whereas Rep52 and Rep40 moderately repress the levels of *PPP1R12C* transcripts (Fig. 7B and C, mCherry). In contrast to the wt Rep proteins, Rep78/Rep68 and Rep52/Rep40 NTP-binding mutants exert different effects on the regulation of *PPP1R12C* promoter activity. Rep78 and Rep68 mutant proteins moderately repress the level of transcription, while the Rep52/Rep40 mutant proteins do not exhibit any negative regulatory effect on *PPP1R12C* promoter activity (Fig. 7B and C, mCherry). Similar observations were made for protein levels (Fig. 7B and C, mCherry).

Having validated that the mechanism by which Rep represses *PPP1R12C* transcription from the bidirectional promoter was similar to what we observed with the PPP1R12Cp-mCherry reporter construct, we next investigated the effects of Rep on the antisense promoter activity. Rep78 and Rep68 inhibit the antisense transcriptional activity by 5- and 7-fold, respectively, while Rep52 and Rep40 repress the antisense promoter by 2- and 3-fold, respectively (Fig. 7B and C, EGFP). Western blot analysis showed strong repression of the antisense promoter by all Rep proteins (Fig. 7B and C, EGFP). In contrast to the wt Rep proteins, all Rep NTP-binding mutants were unable to repress transcription from the antisense promoter (Fig. 7B and C, EGFP). This observation was also made for protein levels (Fig. 7B and C, EGFP).

Altogether, these results highlight the fact that, even though all four Rep proteins inhibit transcription from both promoters, there is a major difference in Rep78/Rep68-mediated transcriptional repression of the sense and antisense promoters. While repression of the *PPP1R12C* promoter requires the NTP-binding motif in the central domain of the Rep proteins, as well as direct interaction with the RBS, only the NTP-binding motif appears to be required for Rep78/Rep68-mediated inhibition of the antisense promoter. With regard to the small Rep proteins, our results suggest that a similar mechanism is involved in Rep52/Rep40-mediated repression of both *PPP1R12C* and antisense promoter activities and that this repression is dependent on the presence of a functional NTP-binding motif.

DISCUSSION

The AAV Rep proteins are multifunctional proteins with the ability to regulate expression from cellular and viral promoters, including the three AAV promoters p5, p19, and p40 (7). The replication phase of the AAV life cycle is strongly dependent on controlled expression of AAV, as well as helper virus proteins, and Rep's role in this regulation has been well characterized (40). However, much less is known about a potential role for Rep in the regulation of cellular proteins that are involved in the different aspects of the AAV life cycle. In particular, it is not known if Rep has the ability to control the expression of the gene in which it integrates to establish latency. In order to gain insight into a potential additional regulatory role for Rep, we investigated if AAV2 infection and Rep expression influence the *PPP1R12C* promoter activity and compared the effect to Rep-mediated repression of the p5 promoter.

We have shown that AAV2 and adenovirus coinfection and the associated increase in Rep expression lead to a decrease in

of each domain are shown above. The black bar shows the position of the NLS. Below are shown schematic representations of the various large Rep proteins coexpressed with the *PPP1R12C*p and p5 mCherry reporter plasmids. All Rep proteins lack endonuclease activity and contain mutations that avoid expression of the small Rep proteins (M225G). The NTP-binding mutants harbor a mutation at amino acid position 340 (K340H). (B) Northern blot analysis shows *PPP1R12C* (left) and p5 (right) transcription levels in 293T cells cotransfected with an mCherry reporter plasmid and various Rep-expressing plasmids. Rep-mediated repression was analyzed on *PPP1R12C*p and p5 reporter plasmids with unaltered, as well as mutated, RBS sequences that abolish Rep binding. The blots were stripped and hybridized with a β-actin cDNA probe (lower blots). Relative fold repression of mCherry expression is indicated at the bottom of the blots. Western blot analysis shows the corresponding mCherry protein levels resulting from *PPP1R12C* (left) and p5 (right) promoter activity (top). Expression of the large Rep proteins and corresponding NTP-binding mutants was determined by hybridization with the N208 antibody. The blots were also incubated with a β-actin antibody, and relative fold repression was determined (indicated below the blots). Slot blot membranes hybridized to an mCherry probe show similar plasmid uptake for all experimental conditions. (C) Determination of average mCherry expression levels from the *PPP1R12C* (left) and p5 (right) promoters from 3 independent experiments by Northern and Western blot analyses. Average fold repression of mCherry expression is indicated at the bottom of each graph. The error bars indicate SEM.

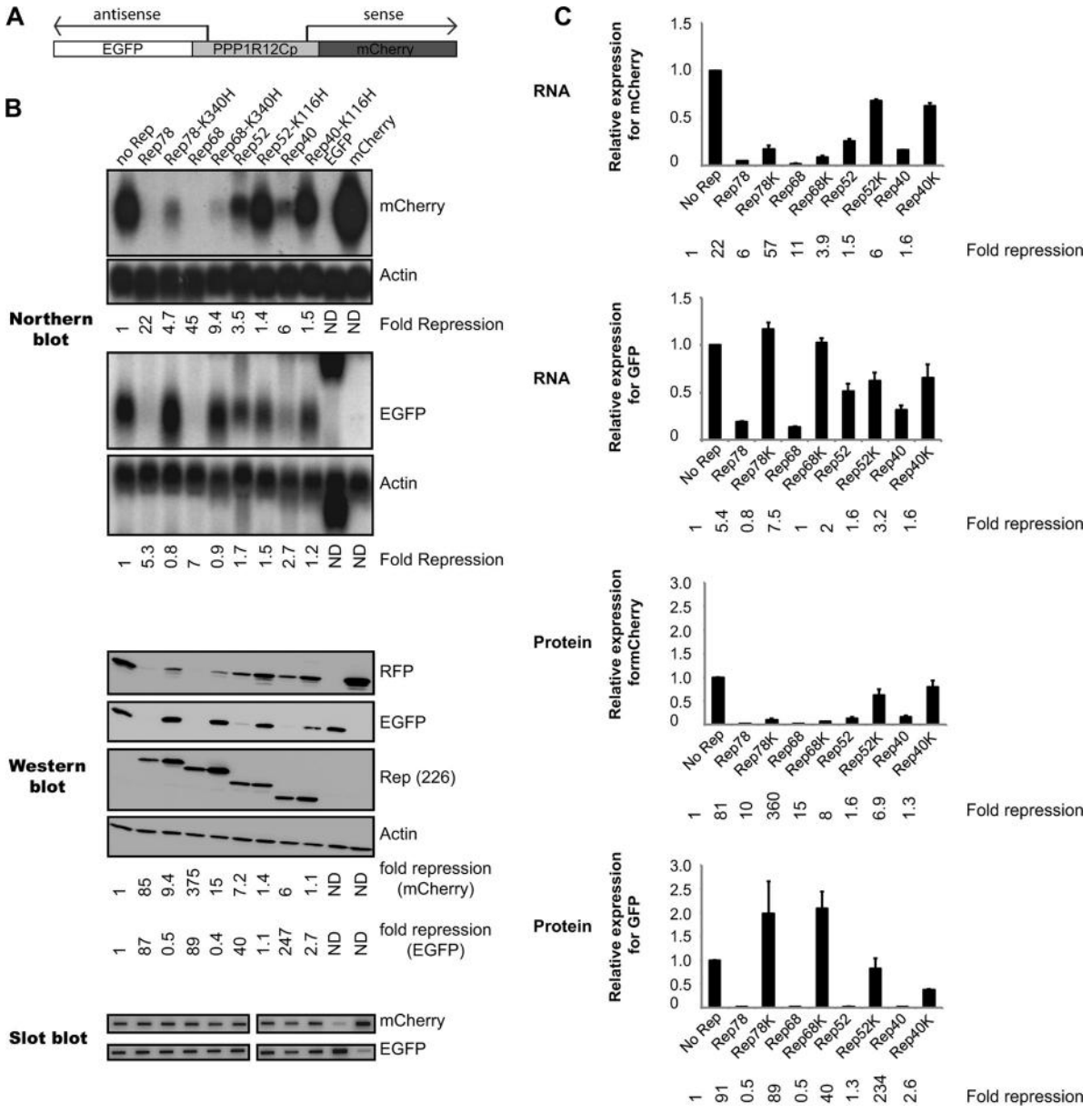


FIG 7 Analysis of the effects of the four Rep proteins and corresponding NTP-binding mutants on gene expression from the *PPP1R12C* sense and antisense promoters. (A) Schematic representation of the *PPP1R12C*p mCherry-GFP reporter plasmid. The Rep proteins that are coexpressed with the reporter plasmid are depicted in Fig. 5 and 6. (B) Northern blot analysis shows mCherry (top) and GFP (bottom) transcripts driven by the sense and antisense promoters, respectively. The blots were stripped and hybridized with a β-actin cDNA probe, and the relative fold repression was calculated (indicated at the bottoms of the mCherry and GFP blots). Western blot analysis shows the corresponding mCherry and GFP protein levels resulting from *PPP1R12C* (top) and antisense promoter activity (second panel from top). Expression of the Rep proteins and corresponding NTP-binding mutants was determined by hybridization with the 226-7 antibody. The blots were also incubated with a β-actin antibody, and the relative fold repression was determined (indicated below the blots). Slot blots were hybridized to an mCherry and a GFP probe to determine cellular uptake of plasmid DNA. ND, not detected. (C) Determination of average EGFP and mCherry expression levels from the *PPP1R12C* promoter from 3 independent experiments by Northern and Western blot analyses. Average fold repression of mCherry and EGFP expression is indicated at the bottom of each graph. The error bars indicate SEM.

PPP1R12C expression levels. The observed infection-induced *PPP1R12C* repression appears to be different in 293T and HeLa cells in that the observed repression can be seen in 293T cells infected with AAV2 at high MOIs, a condition under which Rep transcripts cannot be detected by Northern blotting. However, the presence of E1A and E1B in 293T cells leads to Rep levels that are sufficient to support limited replication in the absence of helper virus (17, 41) and is thus expected to also have an effect on *PPP1R12C* transcription. An additional observation we made in 293T cells is that the level of repression in the presence of adenovirus appears to go up with increasing MOIs despite declining Rep transcripts. This decrease in Rep expression was not observed in HeLa cells and could be due to differences in AAV replication levels and autoregulatory feedback loops (41). The presence of Rep transcripts in the absence of adenovirus infection and down-regulation of Rep expression over time were confirmed in the time

course infection experiments in 293T cells. In summary, we observed that under permissive conditions AAV2 infection causes downregulation of *PPP1R12C* expression, which is potentially mediated by Rep but not necessarily dependent on high levels of Rep transcription.

Transfection-based experiments directly showed that the observed effect is indeed mediated by Rep. In fact, we could show that all four Rep proteins mediate repression of the endogenous gene, as well as the *PPP1R12C* reporter construct, and that this activity is directed by mechanisms similar to those employed by the Rep proteins for the downregulation of the p5 promoter. As was seen for repression of p5 transcription (19, 21), the large Rep proteins display a stronger repressive effect on the *PPP1R12C* promoter reporter constructs than the small Rep proteins (19, 21). This difference could be attributed to the N-terminal DNA binding domain present in Rep78 and Rep68; however, this domain by itself showed a decreased ability to suppress the *PPP1R12C* promoter compared to the full-length large Rep proteins. It is likely that the Rep78/Rep68-induced repression is dependent on the final Rep78/Rep68 oligomeric complex that assembles on the p5 and *PPP1R12C* promoter sites. This event is mediated through specific DNA binding dependent on the presence of the origin binding domain (OBD), interdomain linker, and helicase domain (42). Evidence from structural and biochemical experiments suggests that the assembly of Rep78/Rep68 complexes on RBS-containing DNA sites is both directional and highly cooperative (unpublished data). The X-ray structure of the OBD-RBS complex shows that the Rep molecules bind to the RBS positions in such a way that the C-terminal helicase domain is oriented upstream of the 5′-GCTC-3′ repeats, where it interacts with DNA nonspecifically (43). This model is supported by biochemical data showing that the affinity of Rep68 for DNA with only the minimal RBS site is significantly lower than for sites that also include upstream sequences (32). In addition, footprinting experiments have shown that Rep68 protects regions upstream of the RBS sequence (32). Taken together, it appears that binding of Rep78/Rep68 to RBS DNA sites is a highly cooperative event that requires the participation of all structural domains. Our model proposes that the repression mechanism is a direct effect of Rep78/Rep68 either blocking the start of transcription and/or binding of transcriptional activators in both the p5 and the *PPP1R12C* promoters (Fig. 8). The different directionality of the 5′-GCTC-3′ repeats in the RBS of the p5 versus *PPP1R12C* promoter allows physical interference with the initiation of transcription despite their different positions with respect to the transcriptional start site (Fig. 8). We hypothesize that the K340H mutation may have an effect on DNA affinity or complex formation, explaining the observed lower level of repression in the context of the large Rep proteins and the p5 and sense *PPP1R12C* promoter. However, we cannot exclude the possibility that an additional mechanism that is independent of binding to the RBS but requires interactions with the NTPase domain and cellular proteins also plays a role in Rep78/Rep68-mediated repression.

Our data suggest that this mechanism is likely also responsible for Rep78/Rep68-mediated repression of the antisense promoter, which lacks the RBS and, similar to what was observed for other heterologous promoters (21), requires only an intact NTP-binding motif. The small Rep proteins, which lack the OBD, might exploit the aforementioned interactions with host cell proteins in order to mediate repression of the p5 and *PPP1R12C* sense and antisense promoters (Fig. 5 and 7). In addition, nonspecific DNA binding activities of these proteins may also interfere with the binding activities of proteins from the transcription machinery by masking binding sites; however, further studies need to be performed to answer this question.

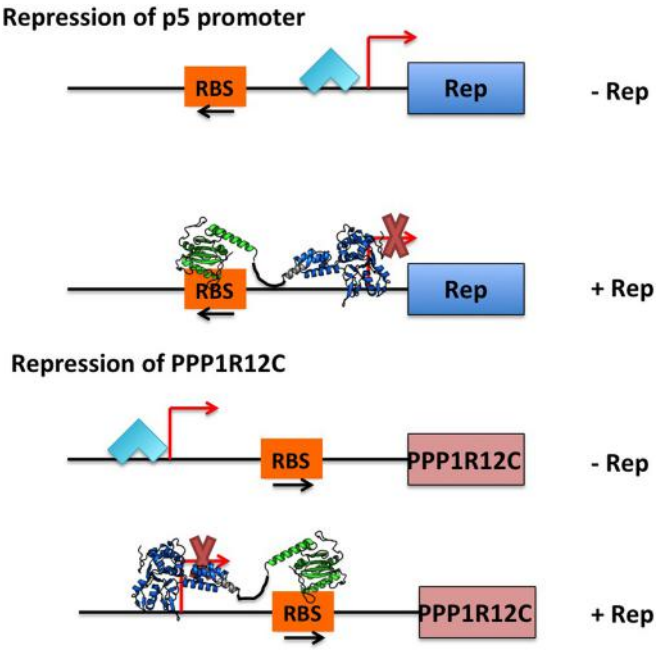


FIG 8 Model of Rep-mediated repression of p5 and integration target site promoters as mediated through OBD-RBS interactions. (Top) Rep78-Rep68-mediated transcriptional repression of the p5 promoter. (Bottom) Rep78-Rep68-mediated transcriptional repression of the *PPP1R12C* sense promoter. The light-blue arrowheads depict transcription factor binding. The red arrows indicate the positions of the transcriptional start sites. The black arrows indicate the directions of the 5′-GCTC-3′ repeats in the RBS. Ribbon representations of the Rep68 OBD (green), interdomain linkers (black), and NTPase domains (blue) depict RBS binding and upstream positioning of the C-terminal SF3 helicase domain, which interferes with the start of transcription and/or binding of transcriptional activators to the promoter.

In summary, in addition to direct interactions with the promoter site, transcriptional repression of both viral p5 and cellular target site promoters may rely on interactions of Rep with factors of the transcription machinery. In fact, DNA-protein interactions, as well as protein-protein interactions, have also been described for Rep-mediated repression of the AdMLP (22). The close proximity of the Rep binding site to the TATA element in AdMLP and the observed interactions of Rep with TATA-binding protein (TBP) have provoked the hypothesis that Rep could repress gene expression by interfering with the formation of the RNA polymerase II preinitiation complex (22).

It could be hypothesized that Rep recruits factors that directly act on RNA polymerase II or exert their function by manipulating the chromatin environment (44). Indeed, several Rep-interacting partners have been identified, among which several proteins are involved in transcriptional regulation: Sp1 (45, 46), high-mobility group 1 (HMG1) nonhistone proteins (47), putative protein kinase PKX and protein kinase A (48, 49), and transcriptional coactivator PC4 (50, 51). The last is an ideal candidate, as Rep-PC4 interactions have been shown to be dependent on Rep's NTPase domain; however, preliminary knockdown experiments in our

laboratory did not affect Rep-mediated repression of the *PPP1R12C* promoter (data not shown).

It is interesting to speculate that AAV2 has adapted to integrate into a chromosomal site, which appears to be regulated by protein complexes that also direct viral gene expression. In fact, once integrated, the virus may exploit these regulatory mechanisms in order to silence the integration site and associated provirus to maintain latency. Furthermore, the observation that the 5′ end of the provirus, containing the promoter region of recombinant or wt AAV2, is usually found in the 5′-3′ transcriptional direction of the *PPP1R12C* gene (52) indicates that these transcriptional protein complexes may also be involved in the formation and positioning of the preintegration complex. Future work directed at identifying Rep's binding partners involved in transcriptional regulation might therefore also shed light on the intricacies of a unique viral integration mechanism.

ACKNOWLEDGMENTS

We thank Chad Swanson and the members of the laboratory for discussions and critical readings of the manuscript.

This work was supported by United Kingdom Medical Research Council grant 1001764. C.R.E. is supported by NIH grant RO1-GM092854.

REFERENCES

1. **Myers MW, Laughlin CA, Jay FT, Carter BJ.** 1980. Adenovirus helper function for growth of adeno-associated virus: effect of temperature-sensitive mutations in adenovirus early gene region 2. J. Virol. **35:**65–75.
2. **Laughlin CA, Jones N, Carter BJ.** 1982. Effect of deletions in adenovirus early region 1 genes upon replication of adeno-associated virus. J. Virol. **41:**868–876.
3. **Carter BJ, Marcus-Sekura CJ, Laughlin CA, Ketner G.** 1983. Properties of an adenovirus type 2 mutant, Ad2dl807, having a deletion near the right-hand genome terminus: failure to help AAV replication. Virology **126:**505–516. http://dx.doi.org/10.1016/S0042-6822(83)80008-7.
4. **West MH, Trempe JP, Tratschin JD, Carter BJ.** 1987. Gene expression in adeno-associated virus vectors: the effects of chimeric mRNA structure, helper virus, and adenovirus VA1 RNA. Virology **160:**38–47. http://dx.doi.org/10.1016/0042-6822(87)90041-9.
5. **Kotin RM, Siniscalco M, Samulski RJ, Zhu XD, Hunter L, Laughlin CA, McLaughlin S, Muzyczka N, Rocchi M, Berns KI.** 1990. Site-specific integration by adeno-associated virus. Proc. Natl. Acad. Sci. U. S. A. **87:**2211–2215. http://dx.doi.org/10.1073/pnas.87.6.2211.
6. **Samulski RJ, Zhu X, Xiao X, Brook JD, Housman DE, Epstein N, Hunter LA.** 1991. Targeted integration of adeno-associated virus (AAV) into human chromosome 19. EMBO J. **10:**3941–3950.
7. **Weitzman MD, Linden RM.** 2011. Adeno-associated virus biology. Methods Mol. Biol. **807:**1–23. http://dx.doi.org/10.1007/978-1-61779-370-7_1.
8. **King JA, Dubielzig R, Grimm D, Kleinschmidt JA.** 2001. DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. EMBO J. **20:**3282–3291. http://dx.doi.org/10.1093/emboj/20.12.3282.
9. **Tratschin JD, Tal J, Carter BJ.** 1986. Negative and positive regulation in trans of gene expression from adeno-associated virus vectors in mammalian cells by a viral rep gene product. Mol. Cell. Biol. **6:**2884–2894.
10. **Chang LS, Shi Y, Shenk T.** 1989. Adeno-associated virus P5 promoter contains an adenovirus E1A-inducible element and a binding site for the major late transcription factor. J. Virol. **63:**3479–3488.
11. **Shi Y, Seto E, Chang LS, Shenk T.** 1991. Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein. Cell **67:**377–388. http://dx.doi.org/10.1016/0092-8674(91)90189-6.
12. **Redemann BE, Mendelson E, Carter BJ.** 1989. Adeno-associated virus rep protein synthesis during productive infection. J. Virol. **63:**873–882.
13. **Mouw MB, Pintel DJ.** 2000. Adeno-associated virus RNAs appear in a temporal order and their splicing is stimulated during coinfection with adenovirus. J. Virol. **74:**9878–9888. http://dx.doi.org/10.1128/JVI.74.21.9878-9888.2000.
14. **McCarty DM, Christensen M, Muzyczka N.** 1991. Sequences required for coordinate induction of adeno-associated virus p19 and p40 promoters by Rep protein. J. Virol. **65:**2936–2945.
15. **Labow MA, Hermonat PL, Berns KI.** 1986. Positive and negative autoregulation of the adeno-associated virus type 2 genome. J. Virol. **60:**251–258.
16. **Pereira DJ, McCarty DM, Muzyczka N.** 1997. The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection. J. Virol. **71:**1079–1088.
17. **Trempe JP, Carter BJ.** 1988. Regulation of adeno-associated virus gene expression in 293 cells: control of mRNA abundance and translation. J. Virol. **62:**68–74.
18. **Beaton A, Palumbo P, Berns KI.** 1989. Expression from the adeno-associated virus p5 and p19 promoters is negatively regulated in trans by the rep protein. J. Virol. **63:**4450–4454.
19. **Kyostio SR, Owens RA, Weitzman MD, Antoni BA, Chejanovsky N, Carter BJ.** 1994. Analysis of adeno-associated virus (AAV) wild-type and mutant Rep proteins for their abilities to negatively regulate AAV p5 and p19 mRNA levels. J. Virol. **68:**2947–2957.
20. **Kyostio SR, Wonderling RS, Owens RA.** 1995. Negative regulation of the adeno-associated virus (AAV) P5 promoter involves both the P5 rep binding site and the consensus ATP-binding motif of the AAV Rep68 protein. J. Virol. **69:**6787–6796.
21. **Horer M, Weger S, Butz K, Hoppe-Seyler F, Geisen C, Kleinschmidt JA.** 1995. Mutational analysis of adeno-associated virus Rep protein-mediated inhibition of heterologous and homologous promoters. J. Virol. **69:**5485–5496.
22. **Needham PG, Casper JM, Kalman-Maltese V, Verrill K, Dignam JD, Trempe JP.** 2006. Adeno-associated virus rep protein-mediated inhibition of transcription of the adenovirus major late promoter in vitro. J. Virol. **80:**6207–6217. http://dx.doi.org/10.1128/JVI.00183-06.
23. **Hermonat PL.** 1991. Inhibition of H-ras expression by the adeno-associated virus Rep78 transformation suppressor gene product. Cancer Res. **51:**3373–3377.
24. **Hermonat PL.** 1994. Down-regulation of the human c-fos and c-myc proto-oncogene promoters by adeno-associated virus Rep78. Cancer Lett. **81:**129–136. http://dx.doi.org/10.1016/0304-3835(94)90193-7.
25. **Wonderling RS, Owens RA.** 1996. The Rep68 protein of adeno-associated virus type 2 stimulates expression of the platelet-derived growth factor B c-sis proto-oncogene. J. Virol. **70:**4783–4786.
26. **Tan I, Ng CH, Lim L, Leung T.** 2001. Phosphorylation of a novel myosin binding subunit of protein phosphatase 1 reveals a conserved mechanism in the regulation of actin cytoskeleton. J. Biol. Chem. **276:**21209–21216. http://dx.doi.org/10.1074/jbc.M102615200.
27. **Dutheil N, Henckaerts E, Kohlbrenner E, Linden RM.** 2009. Transcriptional analysis of the adeno-associated virus integration site. J. Virol. **83:**12512–12525. http://dx.doi.org/10.1128/JVI.01754-09.
28. **Zeltner N, Kohlbrenner E, Clement N, Weber T, Linden RM.** 2010. Near-perfect infectivity of wild-type AAV as benchmark for infectivity of recombinant AAV vectors. Gene Ther. **17:**872–879. http://dx.doi.org/10.1038/gt.2010.27.
29. **Cheung AK, Hoggan MD, Hauswirth WW, Berns KI.** 1980. Integration of the adeno-associated virus genome into cellular DNA in latently infected human Detroit 6 cells. J. Virol. **33:**739–748.
30. **Yoon M, Smith DH, Ward P, Medrano FJ, Aggarwal AK, Linden RM.** 2001. Amino-terminal domain exchange redirects origin-specific interactions of adeno-associated virus rep78 in vitro. J. Virol. **75:**3230–3239. http://dx.doi.org/10.1128/JVI.75.7.3230-3239.2001.
31. **Yoon-Robarts M, Linden RM.** 2003. Identification of active site residues of the adeno-associated virus type 2 Rep endonuclease. J. Biol. Chem. **278:**4912–4918. http://dx.doi.org/10.1074/jbc.M209750200.
32. **Chiorini JA, Wiener SM, Owens RA, Kyostio SR, Kotin RM, Safer B.** 1994. Sequence requirements for stable binding and function of Rep68 on the adeno-associated virus type 2 inverted terminal repeats. J. Virol. **68:**7448–7457.
33. **Chejanovsky N, Carter BJ.** 1990. Mutation of a consensus purine nucleotide binding site in the adeno-associated virus rep gene generates a dominant negative phenotype for DNA replication. J. Virol. **64:**1764–1770.
34. **Schmittgen TD, Livak KJ.** 2008. Analyzing real-time PCR data by the comparative C(T) method. Nat. Protoc. **3:**1101–1108. http://dx.doi.org/10.1038/nprot.2008.73.
35. **Urcelay E, Ward P, Wiener SM, Safer B, Kotin RM.** 1995. Asymmetric

replication in vitro from a human sequence element is dependent on adeno-associated virus Rep protein. J. Virol. **69:**2038–2046.
36. **Weitzman MD, Kyostio SR, Kotin RM, Owens RA.** 1994. Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. Proc. Natl. Acad. Sci. U. S. A. **91:**5808–5812. http://dx.doi.org/10.1073/pnas.91.13.5808.
37. **Smith RH, Kotin RM.** 2000. An adeno-associated virus (AAV) initiator protein, Rep78, catalyzes the cleavage and ligation of single-stranded AAV ori DNA. J. Virol. **74:**3122–3129. http://dx.doi.org/10.1128/JVI.74.7.3122-3129.2000.
38. **Cassell GD, Weitzman MD.** 2004. Characterization of a nuclear localization signal in the C-terminus of the adeno-associated virus Rep68/78 proteins. Virology **327:**206–214. http://dx.doi.org/10.1016/j.virol.2004.06.034.
39. **Wonderling RS, Kyostio SR, Walker SL, Owens RA.** 1997. The Rep68 protein of adeno-associated virus type 2 increases RNA levels from the human cytomegalovirus major immediate early promoter. Virology **236:**167–176. http://dx.doi.org/10.1006/viro.1997.8724.
40. **Weger S, Wistuba A, Grimm D, Kleinschmidt JA.** 1997. Control of adeno-associated virus type 2 cap gene expression: relative influence of helper virus, terminal repeats, and Rep proteins. J. Virol. **71:**8437–8447.
41. **Wang XS, Srivastava A.** 1998. Rescue and autonomous replication of adeno-associated virus type 2 genomes containing Rep-binding site mutations in the viral p5 promoter. J. Virol. **72:**4811–4818.
42. **Zarate-Perez F, Bardelli M, Burgner JW, II, Villamil-Jarauta M, Das K, Kekilli D, Mansilla-Soto J, Linden RM, Escalante CR.** 2012. The interdomain linker of AAV-2 Rep68 is an integral part of its oligomerization domain: role of a conserved SF3 helicase residue in oligomerization. PLoS Pathog. **8:**e1002764. http://dx.doi.org/10.1371/journal.ppat.1002764.
43. **Hickman AB, Ronning DR, Perez ZN, Kotin RM, Dyda F.** 2004. The nuclease domain of adeno-associated virus rep coordinates replication initiation using two distinct DNA recognition interfaces. Mol. Cell **13:**403–414. http://dx.doi.org/10.1016/S1097-2765(04)00023-1.
44. **Saunders A, Core LJ, Lis JT.** 2006. Breaking barriers to transcription elongation. Nat. Rev. Mol. Cell. Biol. **7:**557–567. http://dx.doi.org/10.1038/nrm1981.
45. **Hermonat PL, Santin AD, Batchu RB.** 1996. The adeno-associated virus Rep78 major regulatory/transformation suppressor protein binds cellular Sp1 in vitro and evidence of a biological effect. Cancer Res. **56:**5299–5304.
46. **Pereira DJ, Muzyczka N.** 1997. The adeno-associated virus type 2 p40 promoter requires a proximal Sp1 interaction and a p19 CArG-like element to facilitate Rep transactivation. J. Virol. **71:**4300–4309.
47. **Costello E, Saudan P, Winocour E, Pizer L, Beard P.** 1997. High mobility group chromosomal protein 1 binds to the adeno-associated virus replication protein (Rep) and promotes Rep-mediated site-specific cleavage of DNA, ATPase activity and transcriptional repression. EMBO J. **16:**5943–5954. http://dx.doi.org/10.1093/emboj/16.19.5943.
48. **Chiorini JA, Zimmermann B, Yang L, Smith RH, Ahearn A, Herberg F, Kotin RM.** 1998. Inhibition of PrKX, a novel protein kinase, and the cyclic AMP-dependent protein kinase PKA by the regulatory proteins of adeno-associated virus type 2. Mol. Cell. Biol. **18:**5921–5929.
49. **Di Pasquale G, Stacey SN.** 1998. Adeno-associated virus Rep78 protein interacts with protein kinase A and its homolog PRKX and inhibits CREB-dependent transcriptional activation. J. Virol. **72:**7916–7925.
50. **Weger S, Wendland M, Kleinschmidt JA, Heilbronn R.** 1999. The adeno-associated virus type 2 regulatory proteins rep78 and rep68 interact with the transcriptional coactivator PC4. J. Virol. **73:**260–269.
51. **Nash K, Chen W, Salganik M, Muzyczka N.** 2009. Identification of cellular proteins that interact with the adeno-associated virus rep protein. J. Virol. **83:**454–469. http://dx.doi.org/10.1128/JVI.01939-08.
52. **Henckaerts E, Dutheil N, Zeltner N, Kattman S, Kohlbrenner E, Ward P, Clement N, Rebollo P, Kennedy M, Keller GM, Linden RM.** 2009. Site-specific integration of adeno-associated virus involves partial duplication of the target locus. Proc. Natl. Acad. Sci. U. S. A. **106:**7571–7576. http://dx.doi.org/10.1073/pnas.0806821106.

Epigenetic regulation of adeno-associated virus latency

S. Smith-Moore[1], S. Neil[1], C. Fraefel[2], R.M. Linden[1, 3], H. M. Rowe[4], and E. Henckaerts[1]

Adeno-associated virus (AAV) is a human dependovirus whose low immunogenicity and capacity for long term persistence have led to its widespread use as a vector for gene therapy[1]. Despite great recent successes in AAV-based gene therapy[2–4], further improvements are hindered by an inadequate understanding of various aspects of AAV biology. Of particular importance for the design and safety of AAV vectors, which mimic the latent phase of the viral life cycle, are the mechanisms involved in the regulation of AAV latency. These are poorly understood however. Here we show that recruitment of the corepressor KAP1 to the latent AAV2 genome leads to trimethylation of AAV2-associated histone 3 lysine 9 (H3K9me3). We show that infection by the AAV helper viruses adenovirus type 5 (Ad5) and herpes simplex virus type 1 (HSV-1) leads to KAP1 degradation, and we demonstrate that antagonism of protein phosphatase 1 (PP1) by the AAV Rep52 and Rep78 proteins further counteracts KAP1-mediated repression by enhancing nuclear levels of phosphorylated KAP1-S824. Moreover, we show that this interaction is essential for AAV transcription and replication. This work challenges the currently held model for AAV latency, and introduces not only a new viral mechanism for the counteraction of KAP1 repression, but also the notion that KAP1 targeting may represent a conserved requirement for replication among DNA viruses.

[1]Department of Infectious Diseases, King's College London, Great Maze Pond, London, SE1 9RT, UK. [2]Institute of Virology, University of Zurich, Rämistrasse 71, 8006 Zurich, Switzerland. [3]Present address: Genetic Medicine Institute, Pfizer Inc., 2 Royal College St, London NW1 0NH, UK. [4]Centre for Medical Molecular Virology, Division of Infection and Immunity, University College London, 90 Gower Street, London WC1E 6BT

The rapidly expanding field of viral epigenetics has revealed an intricate web of interactions between viral and host cellular machinery necessary in the establishment of viral latency[5]. In line with the long-standing view of AAV as an anomaly amongst DNA viruses, little is yet known about the nature and contribution of epigenetic marks to the genome organization and temporal gene regulation of AAV. Being a nuclear DNA virus however, it seems unlikely that AAV does not also encounter and interact with similar cellular defenses as other DNA viruses. In this study, we sought to gain insight into the potential role for chromatin structure in the regulation of AAV2 latency by using the screening method BioID[6] to identify new interaction partners for the AAV2 Rep proteins. This led to the identification of Krüppel-associated box domain-associated protein 1 (KAP1/TRIM28/ TIF1-β) (Extended Data Table 1,), a corepressor which acts to form transcriptionally repressive heterochromatin through the recruitment of chromatin-modifying proteins, such as the NuRD histone deacetylase complex and the histone methyltransferase SETDB1[7,8], and which has been linked to the regulation of several viral elements[9–12]. Physical interaction of KAP1 with Rep40, Rep52, and Rep78 was confirmed by biotinylation and immunoprecipitation assays (Extended Data Fig. 1a, b, c, d). To explore the possible significance of the Rep-KAP1 interaction in the AAV life cycle, we performed replication experiments in cells depleted for KAP1. AAV2 genome replication, transcription, and protein expression were enhanced in KAP1-depleted cells coinfected with AAV2 and the helper virus Ad5, but not in cells infected with AAV2 alone (Fig. 1a, b, Extended Data Fig. 2a, b), reflecting AAV2's dependency on helper virus factors to initiate replication and suggesting that KAP1 provides a second layer of regulation, the antagonism of which is necessary but not sufficient for reactivation.

We next asked if KAP1 could be repressing AAV2 through binding of the viral genome and subsequent formation of heterochromatin. We performed KAP1-specific chromatin immunoprecipitation (ChIP) on AAV2-infected 293T cells and analyzed the purified chromatin 2 days after infection by qPCR using primers specific for various areas of the AAV2 genome, as well as *GAPDH* as a negative control, and two zinc finger genes, *ZNF180* and *ZNF274*, as positive controls. KAP1 binding was detected across the *rep* gene, particularly at the 5' and middle regions (Fig. 1c). To determine the functional significance of this binding, we performed ChIP-qPCR for trimethylated H3K9 (H3K9me3), a known marker for KAP1-mediated repression. H3K9me3 was detected across the AAV2 genome (Fig. 1d), and, importantly, histone methylation was lost in KAP1-depleted cells at a ratio similar to that observed for *ZNF180* and *ZNF274*, supporting the notion that KAP1 not only binds the AAV2 genome but also does so in the capacity of a repressor. Importantly, depletion of CHD3 and SETDB1, two members of the KAP1 repressive complex[7,8], also led to an enhancement in AAV2 replication and protein expression (Extended data Fig. 3a, b, c).

Upon DNA damage, ATM-dependent phosphorylation of KAP1 at serine 824 (p-KAP1-S824) results in release of the repressive complex, relaxation of heterochromatin, and relief of transcriptional repression[13,14]. We questioned whether AAV2 replication was associated with the phosphorylation of KAP1-S824, which would suggest a requirement for the inactivation of KAP1 corepressor activity. A dose-dependent increase in p-KAP1-S824 was observed in 293T cells coinfected with Ad5 and increasing concentrations of AAV2, but not a recombinant AAV2 vector (Fig. 2a), suggesting that active replication is necessary to trigger phosphorylation. We next performed a time course to determine if phosphorylation of KAP1-S824 could be related to replication and coinciding with Rep expression (Fig. 2b). In cells coinfected

with Ad5 and AAV2, high levels of p-KAP1-S824 were apparent by 18h post infection, correlating well with the onset of Rep expression and suggesting that the Rep proteins could be directly modulating KAP1 activity.

The large Rep proteins possess endonuclease activity shown to trigger DNA damage[15,16], and all four Rep proteins share a helicase domain with the potential to also disrupt DNA. To determine whether Rep might modulate KAP1 independently from DNA damage, 293T cells expressing various Rep proteins, including endonuclease mutants (Y156F) of Rep68 and Rep78, and catalytic ATPase mutants (K340H) of all four Rep proteins[17,18], were monitored for p-KAP1-S824 levels. Elevated levels of p-KAP1-S824 were apparent in the presence of both Rep52 and Rep78, independently from either endonuclease or ATPase activity (Fig. 2c), suggesting that the Rep proteins mediate phosphorylation of KAP1-S824 via an unknown, DDR-independent pathway. This idea was further supported by transfection and infection experiments performed in the presence of an ATM inhibitor (ATMi), which showed that ATM-activation is not necessary for Rep-mediated phosphorylation of KAP1-S824 (Extended Data Fig. 4a, b, c). Furthermore, expression of a series of C-terminal truncation mutants, in which the C-terminal ZNF domain shared by Rep52 and Rep78 was progressively removed, completely abrogated phosphorylation of KAP1-S824 (Fig. 2d) while having no effect on the Rep-KAP1 interaction (Extended data Fig. 5), suggesting that the Rep proteins act through an intermediary protein(s).

Potential cellular factors that could be interacting with Rep to control the phosphorylation state of KAP1-S824 include protein phosphatase 1 (PP1) and its specific inhibitors. Upon completion of DNA repair, basal levels of p-KAP1-S824 are restored through the combined activities of PP1α/β and protein phosphatase 4 [19,20]. Several regulatory subunits of PP1 were identified as interaction partners for Rep

alongside KAP1 in our original BioID screen, including the nuclear inhibitor of PP1 (NIPP1/PPP1R8), leading us to hypothesize that the Rep proteins could be interfering with this pathway. Using the conserved PP1 consensus binding sequence [KR][X_{0-1}][VI]{P}[FW] as our guideline[21], we discovered one putative binding site in the Rep ATPase domain (K372-W376), partially overlapping with the Walker B motif (Extended data Fig. 6a). Co-IP experiments in cells expressing FLAG-PP1α and Rep52^{K340H} confirmed the physical interaction between Rep and PP1 (Fig. 3a). Use of the repression-deficient ATPase mutant Rep52^{K340H} was necessary to successfully coexpress both proteins at acceptable levels for co-IP. We have shown however that Rep52^{K340H} efficiently phosphorylates KAP1-S824 (Figure 2c). Importantly, mutation of the first lysine in the putative binding site (K372A) impaired this interaction (Fig. 3a) and completely abrogated phosphorylation of KAP1-S824 (Fig. 3b) without affecting Rep ATPase activity (Extended Data Fig. 6b)[18] or interaction with KAP1 (Extended Data Fig. 6c).

Based on our earlier observation that the Rep ZNF domain is also crucial for the phosphorylation of KAP1-S824 (Fig. 2d), we next asked whether this domain could be recruiting the PP1 inhibitor NIPP1, forming a complex with PP1 bound at the ATPase domain. Cross-linked co-IP (CL-IP) experiments between NIPP1 and the various Rep52 ZNF truncation mutants demonstrated that, while NIPP1 readily interacted with Rep52 and Rep52Δ43, this interaction was progressively lost with the two larger truncations, Rep52Δ62 and Rep52Δ87 (Fig. 3c). This displayed a degree of variability however, ranging from a complete loss of binding with both to only a modest decrease as compared to Rep52Δ43. Interestingly, the interaction with NIPP1 was consistently lost with Rep52^{K372A} (Fig. 3c), suggesting that PP1 may in fact represent the main binding partner for Rep and that the Rep ZNF domain acts to

122 stabilize the PP1-NIPP1 holoenzyme, leading to prolonged inhibition of PP1 and
123 increased levels of nuclear p-KAP1-S824. To determine the functional relevance of this
124 interaction in the AAV2 life cycle, we performed replication experiments with an
125 infectious AAV2 plasmid containing the PP1 binding mutation (Rep-K372A). The
126 wild-type plasmid, pAV2 (Rep-WT), was used as a positive control, and the NTP-
127 binding/ATPase mutant K340H (Rep-K340H), which does not support AAV2
128 replication[17], was used as a negative control. While Rep-WT fully supported both
129 replication and transcription in the presence of Ad5, no response was detectable by
130 Rep-K372A (Fig. 3d, e), supporting a model in which Rep-mediated inhibition of PP1
131 is necessary for both DNA replication and transcription, likely through the associated
132 increased levels of p-KAP1-S824.

133 Given that basal levels of Rep expression during latency are not sufficient to
134 counteract KAP1 and that depletion of KAP1 alone is not sufficient to trigger AAV2
135 transcription and replication, we hypothesized that AAV2 helper viruses might act as a
136 biological switch necessary to allow for the upregulation of *rep* expression prior to the
137 onset of KAP1 phosphorylation. KAP1 protein levels were significantly depleted in
138 293T and HeLa cells infected with increasing concentrations of Ad5 (Fig. 4a, b) and
139 were restored in the presence of 5μM of the proteasome inhibitor MG132 (Fig. 4c),
140 suggesting that Ad5 targets KAP1 for degradation. To determine whether this
141 observation extended to other AAV2 helper viruses, we repeated these infection
142 experiments using HSV-1. While the effect was less pronounced, HSV-1 infection also
143 resulted in a 60% depletion of KAP1 over the range of MOIs tested for both 293T and
144 HeLa cells (Fig. 4d, e), suggesting that KAP1 targeting may represent an unknown
145 helper function for AAV2 replication necessary to release the viral genome from its
146 latent state (Extended Data Fig. 7).

This work demonstrates for the first time that AAV2 latency is regulated through the epigenetic modification of its genome, which challenges the long-standing model for AAV latency, whereby AAV is silenced alone through binding of the p5 promoter by cellular factors YY1 and MLTF, and the Rep proteins[22,23]. Given that the biology of AAV vectors likely mimics that of latent viral genomes, the findings presented here have additional relevance for gene therapy. Our data highlight that current production helper plasmids may lack helper genes that may be critical for navigating host responses. Understanding the epigenetic control of AAV may also shed light on the intriguing and unexplained resistance of AAV gene therapy vectors to host shut off, and will undoubtedly contribute to understanding the consequences of integrating wild type and recombinant viruses. In addition, the recent controversial discovery of AAV2 sequences in human liver tumors[24] has caused some to call into question the safety of rAAV vectors and has highlighted the need to further explore AAV2 regulation of latency. These findings may thus provide key insights into the impact and contribution of AAV2 latency on the development of human diseases.

Methods

Cell lines and viruses

293T human embryonic kidney cells and HeLa human cervical epithelial cells were obtained from the American Tissue Culture Collection (ATCC). Cells were cultured in Dulbecco's modified Eagle's (DMEM) (Invitrogen) supplemented with 10% fetal bovine serum (FBS) (Invitrogen) plus 1% Pen/Strep (Sigma) and were tested for mycoplasma once per month.

AAV2 and human adenovirus type 5 (Ad5) were produced and purified as previously described[25].

Plasmids

pcDNA-mycBirAR118G, pCMV-Rep40 (pND229), pCMV-Rep52 (pND230), pCMV-Rep68 Y156F M225G (pND226), and pCMV-Rep78 Y156F M225G (pND227) have been previously described[18]. BirA* and each of the Rep sequences were amplified by PCR, after which overlapping PCR was used to fuse the BirA* fragment to the N-terminus of each Rep fragment. The resulting amplicons were cloned into pcDNA3.1+. FLAG- and T7-tagged Rep proteins were generated by cloning of Rep PCR products into either the pEGFP-C1 vector (Clontech) containing N-terminal FLAG tag or T7 tag, respectively. K372A mutants were generated by site-directed mutagenesis. ZNF truncation mutants were generated by PCR amplification of aa 1-529 (Δ91/87), 1-558 (Δ63), or 1-577 (Δ44), using either pND230 or pND227 as a template. The amplified fragments were then cloned into the N-terminal T7 vector described above. FLAG-PP1α was generated by cloning of PCR-amplified PP1α from an EST clone obtained from Genome Cube (Clone IRAUp969F0817D) into the N-terminal FLAG-vector described above. GFP-NIPP1 was obtained from Angus Lamond [26], through Addgene (#44221).

pC1-FLAG-wtKAP1 was a provided by Helen Rowe and was used for cloning of PCR-amplified wtKAP1 into untagged pcDNA3.1+.

BioID Screening

Ten 10cm dishes of 293T cells per BirA*-Rep construct were transfected using 6µg DNA and 50µl PEI in 500µl serum-free (SF) medium per dish. 6h post-transfection, the medium was replaced with fresh DMEM + 10% FBS, and D-Biotin (Life Technologies) was added to a final concentration of 100µM. 48h post-transfection, cells were harvested for LC-MS/MS analysis as previously described[6]. Mass spectrometry was performed by the KCL proteomics facility at Denmark Hill.

Immunoprecipitation

293T cells were transfected in a 6-well format with 200ng of Rep-expressing constructs, and 250ng of FLAG-PP1α or FAG-GFP using 8ul PEI in 80µl serum free (SF) medium. 48h after transfection, cells were lysed in RIPA buffer, and lysates were incubated with 2µg anti-FLAG (Sigma, F7425) for 1.5h on a rotator at 4°C. 40µl of protein G agarose beads (Pierce) were added and incubated a further 3h. Beads were washed 4 times in RIPA buffer, and proteins were eluted from beads by boiling at 95^{o}C for 10 minutes in 60µl 2X Laemmli buffer.

Cross-linking Immunoprecipitation

293T cells were transfected in a 6-well format with 200ng of Rep-expressing constructs, and 750ng of FLAG-GFP, FLAG-KAP1, or GFP-NIPP1 using using 8ul PEI in 80µl SF medium. 48 hours after transfection cells were fixed in 350ul 0.05% formaldehyde for 10 minutes at 37°C and then quenched in 350ul 0.125M glycine, pH 7, for 5 minutes at room

temperature before being lysed in 500µl cross-linking IP buffer (150mM NaCl, 10mM HEPES pH 7, 6mM $MgCl_2$, 2mM DTT, 10% glycerol, 1X protease inhibitors, 200uM sodium orthovanadate) on ice for 10 minutes. Lysates were subjected to three 10-second cycles of sonication, output ~2, and clarified at 1000 x g for 10 minutes at 4°C. 40ml protein G agarose beads (Pierce) per sample were incubated with 2µg FLAG antibody (Sigma, F7425) for 1.5h on a rotator at 4°C before being added to the cell lysates and incubated a further 3-4 hours. Beads were harvested and washed 4 times in RIPA buffer, and cross-links were reversed in 25 µl reverse cross-link buffer (10mM EDTA, 5mM DTT, 1% SDS) at 65°C for 45 minutes. Proteins were eluted from beads by adding 3µl of 2X Laemmli SDS buffer and boiling at 95^{o}C for 10 minutes.

Western Blot Analysis

Cells were lysed in RIPA buffer, and proteins were separated on a 6-12% SDS-PAGE gel and transferred to a nitrocellulose membrane (Hybond-C Extra nitrocellulose, Amersham Biosciences). Membranes were blocked with either 5% nonfat dry milk or 2.5% BSA (for phospho-antibodies) in PBS containing 0.5% Tween-20 (PBST) for 45 minutes at RT and then incubated with primary antibody for 2h at room temperature. The membranes were then washed 3x10' in PBST and incubated with HRP-conjugated anti-mouse or anti-rabbit IgG (BioRad) for 1h at RT. After 3x10' washes in PBST, membranes were developed using West Pico ECL reagent (Thermo Scientific). The following primary antibodies were used**:** HSP90 (Santa Cruz, sc-69703; 1:10,000), KAP1 (Chemicon, MAB3662; 1:000), p-KAP1-S824 (Bethyl, A300-767A; 1:2000), VP (ARP, 03-61058; 1:500), Rep (Progen, 61069; 1:100), dsRed (Clontech, 632496; 1:2000), CHD3 (Bethyl, A301-219A; 1:4000), SETDB1 (Abcam, ab12317; 1:1000), FLAG

238 (Sigma, F1804; 1:1000), T7 (Merck Millipore, 69522-3; 1:10,000), pChk2 (NEB, 2661S;
239 1:1000), GFP (Roche, 11814460001; 1;5000) and avidin peroxidase (Sigma; 1:8000).
240
241 **Lentiviral transductions**
242 pC-SIREN-based lentiviral vectors expressing either a hairpin targeting the 3'UTR of
243 KAP1(shKAP1; GATCCGCCTGGCTCTGTTCTCTGTCCTTTCAAGAGAAGGA
244 CAGAGAACAGAG CCAGGTTTTTTACGCGTG) or the corresponding empty vector
245 (shEMPTY) were provided by Helen Rowe. Lentiviral vector-containing supernatants
246 were produced by the common triple transfection method using VSV-G, HIV-Gag, and
247 the lentiviral vector in a 3:2:1 molar ratio. Supernatants were harvested 48 and 72 hours
248 after transfection, pooled, filtered, and frozen at -80^{o}C until use. For knockdown, 1 x 10^{6}
249 293T cells were transduced in a 6-well format using 1.2-1.6 mL of either shKAP1 or
250 shEMPTY diluted with the appropriate amount of DMEM + 10% FBS. 24h after
251 transduction, cells were re-plated at a density of 5 x 10^{5} cells/mL for infection with
252 AAV2/Ad5 the next day.
253
254 **siRNA transfections.** In a 24-well format, 2 x 10^{5} cells were transfected with 50nM
255 (siKAP1.2 and siKAP1.4) or 100nM (siCHD3 and siSETDB1) siRNA using 2ul
256 Dharmafect (Dharmacon) in 50ul Optim-mem (Gibco). 6h later, medium was replaced
257 with fresh DMEM + 10% FBS. 24h after transfection, cells were re-plated into 12-well
258 format. 36h after transfection, cells were subjected to a second transfection as described
259 above, using 4ul Dharmafect in 100ul Opti-mem. 4h after the second transfection, cells
260 were infected with 10 IU/cell AAV2 and 2 PFU/cell of Ad5 as described for the viral
261 replication experiments.
262

Infections

For KAP1 depletion, 293T cells were transduced with a lentiviral vector expressing either a hairpin targeting the 3'UTR of KAP1, or the corresponding empty vector, 48h before infection with AAV2/Ad5. Cells were infected at ~80% confluency by adding AAV2 at the stated MOI in ~2/5 the normal well volume. 2h after AAV2 infection, Ad5 was added at an MOI of 2 PFU/cell, and medium was replaced with fresh DMEM + 10% FBS 1h after Ad5 infection. Cells were harvested for qPCR, RT-qPCR, or western blot ~42h after infection, or when they displayed optimal cytopathic effect (CPE). Optimal CPE is defined by cells that display a rounded and enlarged phenotype, as opposed to the normal "star-shaped" morphology of HEK293T cells, and which are beginning to detach but still appear bright and healthy. Cells that had completely lifted by the time of harvest were deemed too advanced in the infection cycle and were excluded from analysis.

ChIP-qPCR

Cells were cross-linked in their medium in 1% formaldehyde (10' at room temperature) and quenched with 0.125M glycine (5' at room temperature) before being lysed in 1mL/1x10^8 cells lysis buffer (50mM Tris-HCl, pH 8, 10mM EDTA, 1% SDS, 1x protease inhibitors) for 10' on ice. Lysates were sonicated to obtain 200- to 500-bp fragments (15 x 30" cycles with 90" intervals, output ~2). 10ml of lysates were used to assess sonication efficiency by reverse cross-linking for 15' at 95°C and then incubating with RNAse A for 30' at 37 degrees. DNA was extracted and visualized on a 1.5% agarose gel. The remaining lysates were clarified at 13,000 rpm for 10' at 4 degrees. The equivalent of 2x10^6 cells was diluted 25-fold in RIPA buffer (50mM Tris pH8, 150mM NaCl, 2mM EDTA, pH 8, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 1x protease inhibitors) and pre-cleared with 80ml protein G agarose beads (pre-blocked in 0.1mg/mL

BSA for 30') for 2h on a rotator at 4°C. For the immunoprecipitation, antibodies were added to lysates and incubated with antibody for 1h on a rotator at 4°C (5mg IgG [Abcam; ab37415], 4mg H3K9me3 [Abcam; ab8893], 1mg KAP1 [Abcam; ab10483]), before adding 80ml pre-blocked beads and incubating overnight as above. Beads were harvested and washed 4 times in RIPA buffer, 4 times in high salt wash (20mM Tris-HCl, pH 8, 1mM EDTA, 500mM NaCl, 0.5% NP-40, 1x protease inhibitors), 4 times in TE buffer (10mM Tris-HCl, pH8, 1mM EDTA), and eluted in 160ml elution buffer (100mM $HaHCO_3$, 1% SDS) for 15' at 30°C. Cross-links were reversed by adding NaCl to a final concentration of 0.2M and incubating overnight at 67°C. Eluates were then incubated with 2ml RNase A (10mg/mL) and 2ml proteinase K (20mg/mL) at 45 degrees for 1h. DNA was extracted using a PCR purification kit (Qiagen) and analyzed by qPCR using primers specific for GAPDH, ZNF180, ZNF274, or various regions of the AAV2 genome. Purified chromatin was diluted 10-fold and quantified by real-time PCR using the SYBR Green JumpStart Taq ReadyMix for QPCR (Sigma-Aldrich) using an ABI PRISM system (Applied Biosystems). Primer sequences are listed in the Extended Data Table 2. CT values for "10% input" were adjusted by subtracting 3.322 cycles to correct for the 10-fold dilution factor (https://www.thermofisher.com/uk/en/home/life-science/epigenetics-noncoding-rna-research/chromatin-remodeling/chromatin-immunoprecipitation-chip/chip-analysis.html). Percent input was then calculated as follows: 100 x 2^-(CT of adjusted 10% input - CT of ChIP-ed DNA). Percent input for each antibody was then normalized to values for IgG to calculate final fold enrichment.

Real-time PCR

For analysis of viral replication, total DNA was extracted using the Qiagen DNAeasy Blood and Tissue DNA extraction kit. Viral DNA was quantified by real-time PCR using

313 the SYBR Green JumpStart Taq ReadyMix for qPCR (Sigma-Aldrich) using an ABI
314 PRISM system (Applied Biosystems). Cap and Ad5 100kd-specific primers and a pDG-
315 based[27] standard curve were used for absolute quantification; the signal was normalised
316 to cyclophilin. Primers: Cap FW (5' – TTCTCAGATGCTGC GTACCGGAAA – 3'),
317 Cap RV (5' – TCTGCCATTGAGGTGGTACTTGGT – 3'), Ad5 100kd FW (5'-
318 TCATTACCCAGGGCCACATT – 3'), Ad5 100kd RV (5' –
319 CCTCGTCCAAAACCTCCTCT – 3'), cyclophilin FW (5' – TGCTGGACCCAAC
320 ACAAATG – 3'), cyclophilin RV (5' – TGCCATCCAACCACTCAGTCT – 3').
321
322 **qRT-PCR**
323 Total RNA was extracted using the RNeasy kit (Qiagen) after DNAseI (Qiagen)
324 treatment for 15 minutes at 37˚C. Reverse transcription was performed using the High
325 Capacity Reverse Transcription kit (Applied Biosystems). cDNA was quantified by real-
326 time qPCR on an ABI PRISM system (Applied Biosystems) using the TaqMan Universal
327 PCR master mix (Life Technologies and custom designed primer-probe mixes (Eurofins).
328 Primers: p5 FW (5'- AACAAGGTGGTGGATGAGT - 3'), p5 RV (5' –
329 CGTTTACGCTCCGTGAGATT - 3'), p40 FW (5' – GGAAGCAAGGCTCAGAGAAA
330 -3') and p40 RV (5' – CCTCTCTGGAGGTTGGTAGATA - 3'). Probes: p5 (5' - FAM-
331 ACGTGGTTGAGGTGGAGCATGAT-TAM - 3'), and p40 (5' - FAM-
332 AGGAAATCAGGACAA CCAATCCCGT-TAM - 3'). Relative expression levels were
333 determined with the ΔΔCt quantification method using 18s ribosomal RNA (Taqman Pre-
334 developed assay reagents, human 18S rRNA, Applied Biosystems) as a housekeeping
335 reference gene.
336
337

Analysis of p-KAP1-S824 levels

Phosphorylation of KAP1-S824 was investigated in 293T cells that were either infected with AAV2/Ad5, or transfected with various Rep-expressing constructs using linear PEI. Where relevant, cells were pretreated with either DMSO or 10μM ATMi 4h prior to infection/trasnsfection, and inhibitors were maintained throughout. Infections were performed as described above, and transfections were performed at ~70% confluency using 1μg DNA/$8x10^5$ cells and 4 μl PEI/1μg DNA. Medium was changed 6h after transfection, and cells were harvested for western blot 27h after infection/transfection.

Immunofluorescence

293T cells were seeded at a density of $1x10^5$ on poly-L-lysine (Sigma) coated coverslips in 24-well plates the day prior to transfection. 4h prior to transfection, DMSO or ATMi was added to the appropriate wells to a final concentration of 10μM. Cells were then transfected with 20ng of empty vector or pRep78-GFP using 2μl Lipofectamine 2000 in 50μl serum-free Opti-mem. The next day, cells were infected with Ad5 (2 PFU/cell) in a total volume of 160μl for 1 h, after which the medium was replaced for fresh DMEM + 10% FBS. Cells were fixed 24h after Ad5 infection in 4% PFA for 10' at room temperature, washed in PBS, permeabilized in 0.1% Triton-X-100 for 10' at room temperature, and washed again in PBS. Cells were then incubated with primary antibody (α-p-KAP1-S824 antibody; 1:1000) diluted in PBS + 1% BSA for 2h at room temperature, washed, and then incubated with secondary antibody (Biolegend; rabbit IgG2b-AlexaFluor 594, 1 μg/ml (1:1000) diluted in PBS for 1h at room temperature. Cells were then washed a final time and mounted in Prolong Gold Antifade Reagent (Invitrogen). Images were visualized using an Eclipse Ti-E Inverted confocal microscope and analyzed with NIS Elements C software.

Analysis of Rep-K372A replication efficiency. 8 x 10^6 293T cells in a 10 cm dish were transfected with 16ug of pAV2-Rep-WT, pAV2-Rep-K340H, or pAV2-Rep-K372A using 90µl of PEI and 500µl SF medium. 4h later, cells were infected with 2 PFU/cell of Ad5, and the medium was replaced 2h after infection. Cells were harvested for DNA, RNA, and protein extraction 72h after transfection.

References

1. Hastie, E. & Samulski, R. J. AAV at 50: A golden anniversary of discovery, research, and gene therapy success, a personal perspective. *Hum. Gene Ther.* **265,** 1–24 (2015).
2. Nathwani, A. C. *et al.* Long-Term Safety and Efficacy of Factor IX Gene Therapy in Hemophilia B. *N. Engl. J. Med.* **371,** 1994–2004 (2014).
3. Gaudet, D., Méthot, J. & Kastelein, J. Gene therapy for lipoprotein lipase deficiency. *Curr. Opin. Lipidol.* **23,** 310–20 (2012).
4. Bennet, J. *et al.* Safety and durability of effect of contralateral-eye administration of AAV2 gene therapy in patients with childhood-onset blindness caused by RPE65 mutations: a follow-on phase 1 trial. *Lancet* (2016).
5. Lieberman, P. Epigenetics and Genetics of Viral Latency. *Cell Host Microbe* **19,** 619–628 (2016).
6. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* **196,** 801–810 (2012).
7. Schultz, D. C., Friedman, J. R. & Rauscher 3rd, F. J. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. *Genes Dev* **15,** 428–443 (2001).
8. Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher 3rd, F. J. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev* **16,** 919–932 (2002).
9. Wolf, D. & Goff, S. P. TRIM28 mediates primer binding site-targeted silencing of murine

leukemia virus in embryonic cells. *Cell* **131,** 46–57 (2007).

10. Rowe, H. M. *et al.* KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463,** 237–240 (2010).

11. Chang, P. C. *et al.* Kruppel-associated box domain-associated protein-1 as a latency regulator for Kaposi's sarcoma-associated herpesvirus and its modulation by the viral protein kinase. *Cancer Res* **69,** 5681–5689 (2009).

12. Rauwel, B. *et al.* Release of human cytomegalovirus from latency by a KAP1/TRIM28 phosphorylation switch. *Elife* **4,** (2015).

13. Goodarzi, A. A., Kurka, T. & Jeggo, P. A. KAP-1 phosphorylation regulates CHD3 nucleosome remodeling during the DNA double-strand break response. *Nat Struct Mol Biol* **18,** 831–839 (2011).

14. Li, X. *et al.* Role for KAP1 serine 824 phosphorylation and sumoylation/desumoylation switch in regulating KAP1-mediated transcriptional repression. *J Biol Chem* **282,** 36177–36189 (2007).

15. Berthet, C., Raj, K., Saudan, P. & Beard, P. How adeno-associated virus Rep78 protein arrests cells completely in S phase. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 13634–9 (2005).

16. Schwartz, R. A., Carson, C. T., Schuberth, C. & Weitzman, M. D. Adeno-associated virus replication induces a DNA damage response coordinated by DNA-dependent protein kinase. *J Virol* **83,** 6269–6278 (2009).

17. Chejanovsky, N. & Carter, B. J. Mutation of a consensus purine nucleotide binding site in the adeno-associated virus rep gene generates a dominant negative phenotype for DNA replication. *J Virol* **64,** 1764–1770 (1990).

18. Dutheil, N. *et al.* Adeno-associated virus Rep represses the human integration site promoter by two pathways that are similar to those required for the regulation of the viral p5 promoter. *J Virol* **88,** 8227–8241 (2014).

19. Li, X. *et al.* SUMOylation of the transcriptional co-repressor KAP1 is regulated by the serine and threonine phosphatase PP1. *Sci Signal* **3,** ra32 (2010).

20. Pfeifer, G. P. Protein phosphatase PP4: role in dephosphorylation of KAP1 and DNA strand break repair. *Cell Cycle* **11,** 2590–2591 (2012).

21. Meiselbach, H., Sticht, H. & Enz, R. Structural analysis of the protein phosphatase 1 docking motif: molecular description of binding specificities identifies interacting proteins. *Chem Biol* **13,**

49–59 (2006).

22. Chang, L. S., Shi, Y. & Shenk, T. Adeno-associated virus P5 promoter contains an adenovirus E1A-inducible element and a binding site for the major late transcription factor. *J Virol* **63,** 3479–3488 (1989).

23. Seto, E., Shi, Y. & Shenk, T. YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro. *Nature* **354,** 241–245 (1991).

24. Nault, J.-C. *et al.* Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas. *Nat. Genet.* **47,** 1–15 (2015).

25. Zeltner, N., Kohlbrenner, E., Clement, N., Weber, T. & Linden, R. M. Near-perfect infectivity of wild-type AAV as benchmark for infectivity of recombinant AAV vectors. *Gene Ther* **17,** 872–879 (2010).

26. Trinkle-Mulcahy, L., Sleeman, J. E. & Lamond, a I. Dynamic targeting of protein phosphatase 1 within the nuclei of living mammalian cells. *J. Cell Sci.* **114,** 4219–4228 (2001).

27. Grimm, D., Kern, A., Rittner, K. & Kleinschmidt, J. A. Novel tools for production and purification of recombinant adenoassociated virus vectors. *Hum. Gene Ther.* **9,** 2745–2760 (1998).

28. Hörer, M. *et al.* Mutational analysis of adeno-associated virus Rep protein-mediated inhibition of heterologous and homologous promoters. *J. Virol.* **69,** 5485–96 (1995).

Acknowledgements We thank M. Bardelli, G. Berger, R. Galão, T.Foster, and S. Pickering for their technical assistance. We thank C. Swanson for his insightful comments and support. This work was supported by UK MRC grants 1001764 to R.M.L., and MR/N022890/1 to E. H.

Author Contributions S.S conceived and performed experiments and wrote the manuscript. S.N. and H.M.R. provided reagents, expertise, and feedback. C.F. provided reagents. R.M.L. provided expertise and feedback, and secured funding. E.H. conceived experiments and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to E. H. (els.henckaerts@kcl.ac.uk)

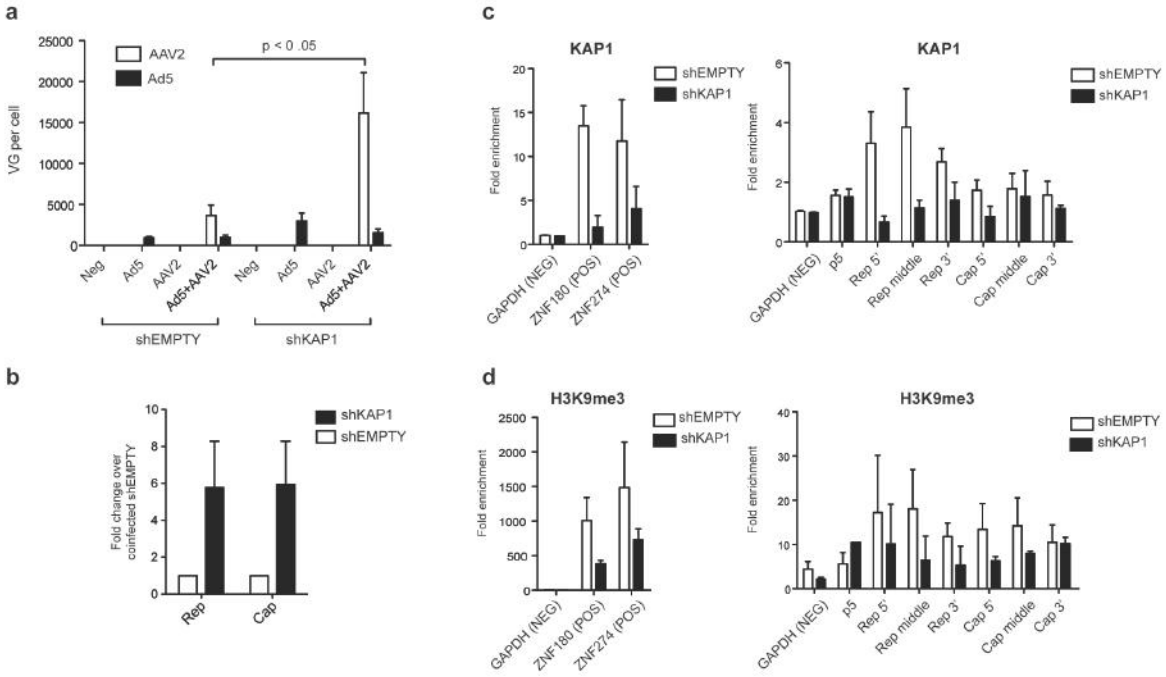


Figure 1. The latent AAV2 genome is methylated by recruitment of KAP1 to the *Rep* ORF

a-b, AAV2 replication in control (shEMPTY) or KAP1-depleted (shKAP1) 293T cells. **a**, Viral genome replication. **b**, *Rep* and *cap* expression. Expression levels represent fold changes over control cells infected with AAV2 + Ad5. Data are reported as mean±SEM, n=5, and statistical significance was determined by unpaired t-test. **c**, ChIP-qPCR performed on AAV2-infected shEMPTY or shKAP1 293T cells using anti-KAP1 antibody or IgG. Purified chromatin was analyzed by qPCR using primers for the viral p5 promoter or various regions of the *rep* and *cap* ORFs (right panel). *GAPDH* was used as a negative control, and the zinc finger genes *ZNF180* and *ZNF274* were used as positive controls (left panel). **d**, ChIP-qPCR performed as described above, using anti-H3K9me3 antibody. Values are reported as mean±SEM for 3 independent experiments.

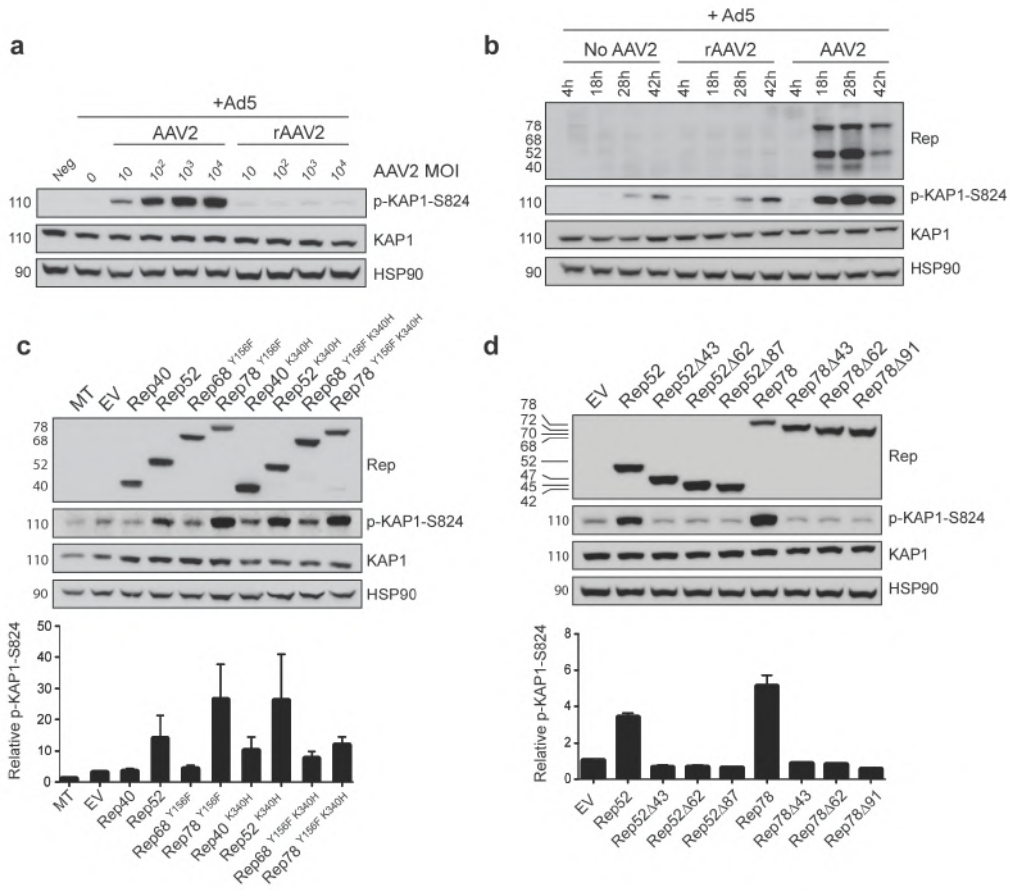


Figure 2. Rep52 and Rep78 lead to phosphorylation of KAP1 in a ZFN domain-dependent manner. **a**, p-KAP1-S824 in 293T cells infected with increasing MOIs of either AAV2 or rAAV2 in the presence of Ad5. **b**, p-KAP1-S824 in 293T cells infected with Ad5 alone, or coinfected with Ad5 and either AAV2 or rAAV2 (MOI = 10^3) monitored at 4, 18, 24, and 42 h post infection. **c**, p-KAP1-S824 in 293T cells expressing various Rep proteins. **d**, p-KAP1-S824 in 293T cells transfected with full length Rep52 and Rep78, or truncation mutants in which the C-terminal ZNF domain is progressively removed. Values are reported as mean±SEM, n=3.

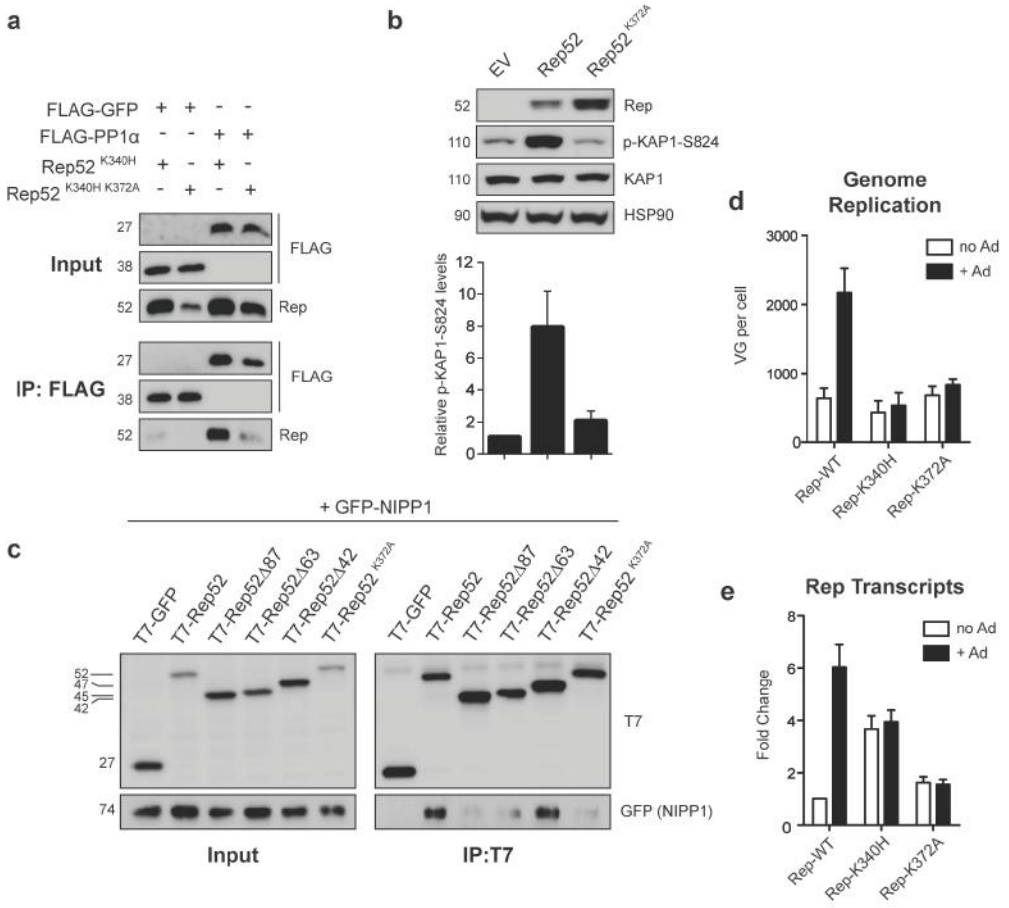


Figure 3. Binding of PP1α and its negative regulator NIPP1 by Rep is necessary for phosphorylation of KAP1 and is required for AAV2 genome replication and transcription.

a, Co-IP of FLAG-tagged proteins from 293T cells expressing FLAG-PP1α or a FLAG-GFP control and the indicated Rep proteins. **b**, Immunoblot of p-KAP1-S824 in 293T cells transfected with either wild-type Rep52, or Rep52-K372A. Protein levels were quantified using ImageJ software. Values are reported as mean±SEM, n=3. **c**, CL-IP of T7-tagged proteins from 293T cells expressing T7-GFP, or the indicated T7-Rep proteins and GFP-NIPP1. All immunoblots are representative of at least 3 independent experiments. **d-e**, 293T cells transfected with pAV2-WT, pAV2-K340H, or pAV2-K372A and then infected with Ad5 were analyzed for AAV2 genome replication by qPCR, **d**, and *rep* and *cap* transcription by RT-qPCR, **e**. Values are reported as mean±SEM, n=3.

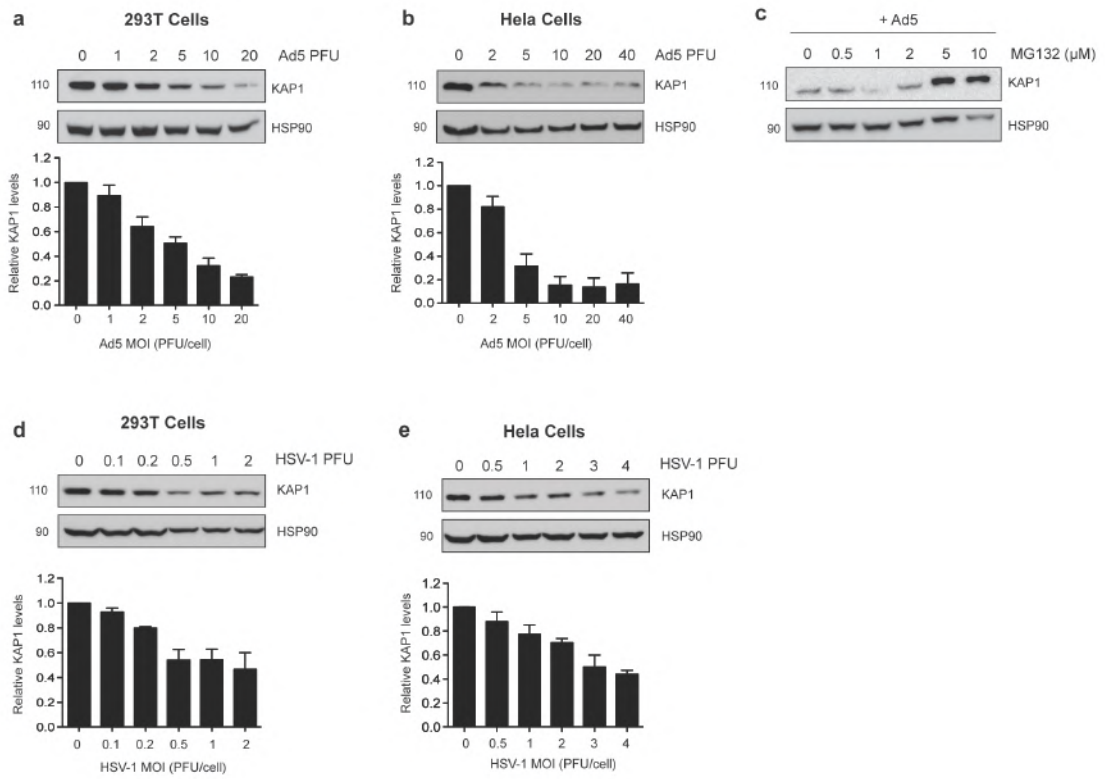


Figure 4. Ad5 and HSV-1 infection leads to KAP1 degradation.

a-b, Immunoblot of KAP1 in 293T cells **a** and Hela cells **b** infected with Ad5 at the stated MOI (PFU/cell). Values are reported as mean±SEM, n=3. **c**, Immunoblot of KAP1 in HeLa cells treated with MG132 at the stated concentrations and infected with 10 PFU/cell of Ad5. **d-e**, Immunoblot of endogenous KAP1 in 293T, **c**, and Hela, **d**, cells infected with HSV-1. Values are reported as mean±SEM, n=3.

505

506 **Extended Data**

507

508 **Extended Data Table 1. Peptides identified by BioID for KAP1 and various known**

509 **interaction partners of the Rep proteins**

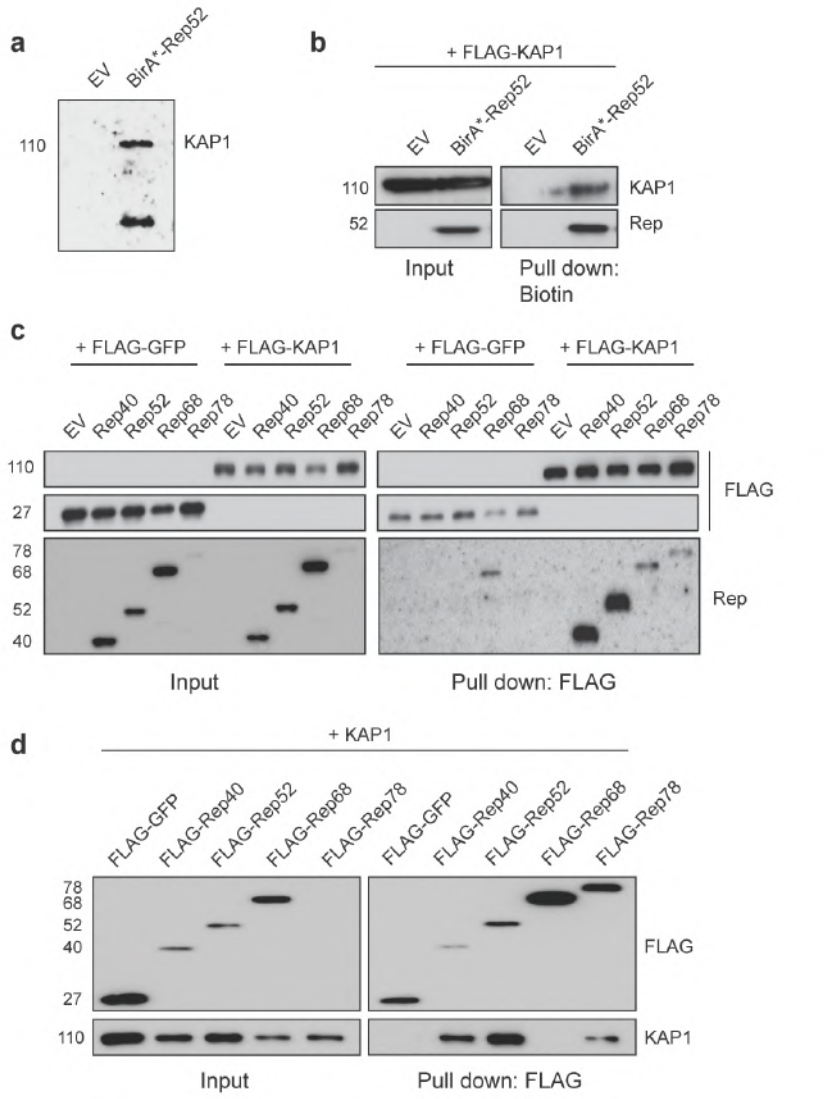
Identified Protein	Accession Number	Bait Protein	Unique Peptides	Sequence Coverage	Protein ID Probability
KAP1	Q13263	BirA*-Rep40	2	3.1%	100%
		BirA*-Rep52	8	12.9%	100%
		BirA*-Rep68	3	6.0%	100%
		BirA*-Rep78	1	1.3%	100%
RUVBL1	Q9Y265	BirA*-Rep40	2	7.2%	100%
		BirA*-Rep52	7	22.1%	100%
		BirA*-Rep68	1	2.9%	99%
		BirA*-Rep78	1	2.9%	99%
MRE11	P49959	BirA*-Rep40	1	2.0%	99%
		BirA*-Rep52	8	19%	100%
SNW1	Q5R7R9	BirA*-Rep40	5	23.7%	100%
		BirA*-Rep52	9	24.1%	100%
		BirA*-Rep68	1	2.8%	100%
		BirA*-Rep78	2	4.3%	100%
MDC1	Q14676	BirA*-Rep40	2	2.0%	100%
		BirA*-Rep52	4	3.3%	100%
		BirA*-Rep68	1	0.5%	100%
TAF1/SET	Q01105	BirA*-Rep40	1	4.8%	100%
		BirA*-Rep68	1	4.8%	100%
NUCLEOLIN	P19338	BirA*-Rep40	1	2%	99%
		BirA*-Rep52	10	16.2%	100%
		BirA*-Rep68	13	18%	100%
		BirA*-Rep78	4	6.2%	100%

510

511

Extended Data Table 2. ChIP-qPCR primers

Gene	FW	RV
GAPDH	CACCGTCAAGGCTGAGAACG	ATACCCAAGGGAGCCACACC
ZNF180	TGATGCACAATAAGTCGAGCA	TGCAGTCAATGTGGGAAGTC
ZNF274	GGAGAAATCCCATGAGGGTAA	GGCTTTTGTGAGAATGTTTTCC
p5	CTGTATTAGAGGTCACGTGAGTG	TCAAACCTCCCGCTTCAAA
Rep 5’	CCGAGAAGGAATGGGAGTT	CCATTCCGTCAGAAAGTCG
Rep middle	GCCTTGGACAATGCGGGAAAGATT	TGTCGACACAGTCGTTGAAGGGAA
Rep 3’	TTCCCGTGTCAGAATCTCAA	CCAAATCCACATTGACCAGA
Cap 5’	GACAGTGGTGGAAGCTCAAA	TTGTACCCAGGAAGCACAAG
Cap middle	TTCTCAGATGCTGCGTACCGGAAA	TCTGCCATTGAGGTGGTACTTGGT
Cap 3’	GTCAGCGTGGAGATCGAGT	AGGCTCTGAATACACGCCAT

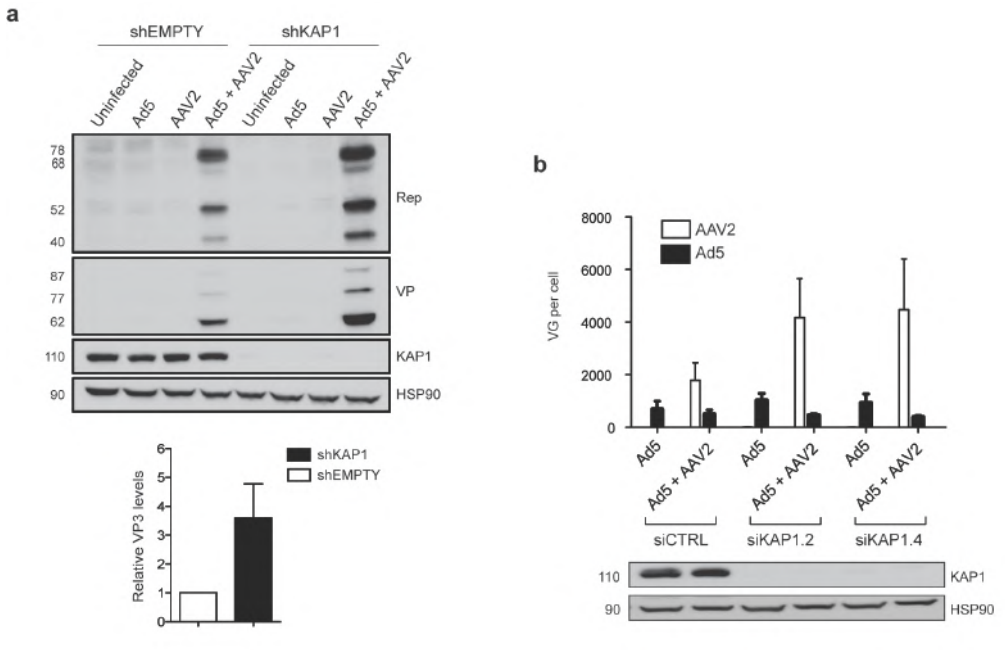


517

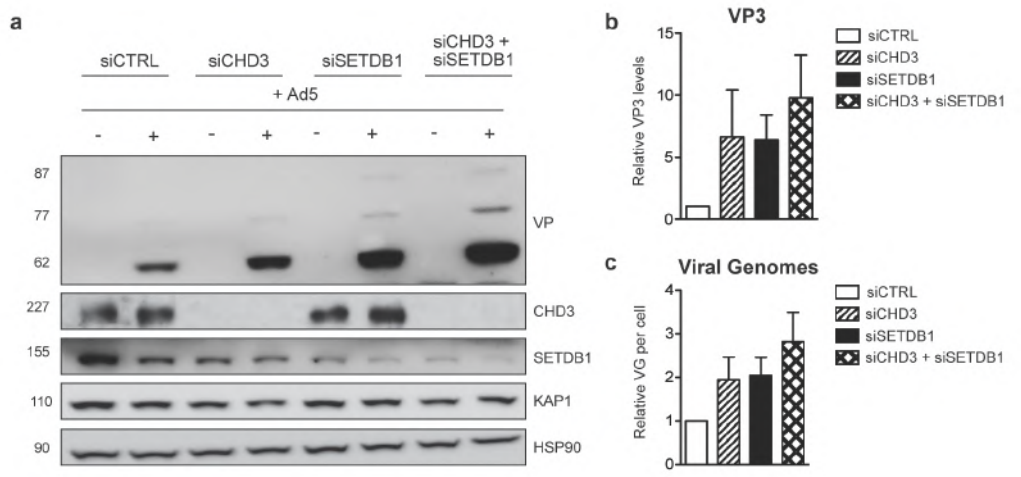
518 **Extended Data Figure 1. The AAV2 Rep Proteins Physically Interact with KAP1**

519 **a**, Immunoblot of biotinylated proteins purified from BirA*-Rep52 BioID screen using anti-
520 KAP1. **b**, Verification of BioID using exogenous FLAG-KAP1; purified biotinylated proteins
521 from 293T cells expressing FLAG-KAP1 with either empty vector (EV) or BirA*-Rep52 were
522 analyzed for Rep and KAP1 by western blot. **c**, Cross-linked co-IP for FLAG-tagged proteins
523 from 293T cells expressing FLAG-KAP1 or a FLAG-GFP control and each of the four Rep
524 proteins. **d**, Cross-linked co-IP for FLAG-tagged proteins from lysates of 293T cells expressing
525 FLAG-GFP, FLAG-Rep40, FLAG-Rep52, FLAG-Rep68, or FLAG-Rep78 and KAP1 from
526 transfected 293T cells.

527

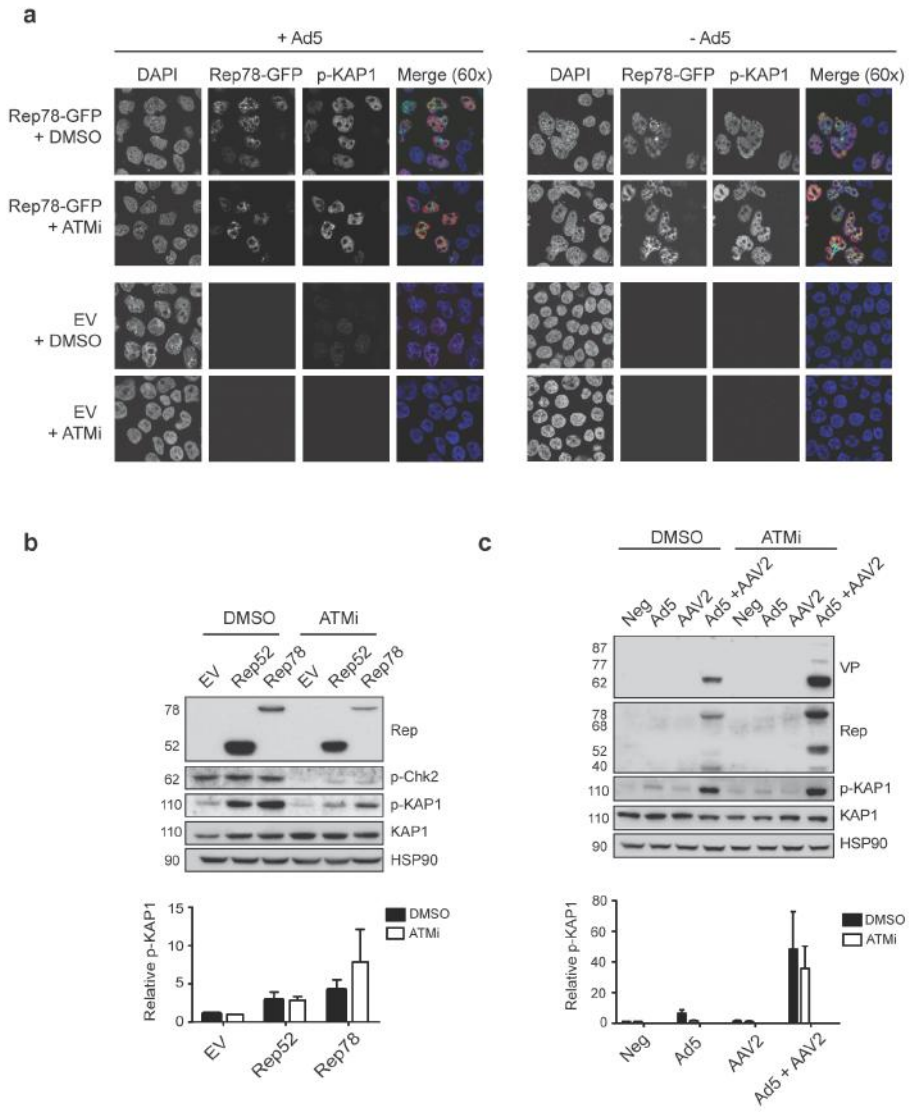


Extended Data Figure 2. AAV2 replication and protein expression in KAP1-depleted cells. **a**, viral protein (VP) expression and KAP1 knockdown efficiency in AAV2 and Ad5 infected control (shEMPTY) and KAP1-depleted (shKAP1) 293T cells. Data are reported as mean±SEM, n=3 **b**, AAV2 replication in control (siCTRL) or KAP1-depleted (siKAP1.2/siKAP1.4) cells. Viral genome replication was analyzed by qPCR, and KAP1 knockdown efficiency was analyzed by western blot. Data are reported as mean±SEM, n=4.



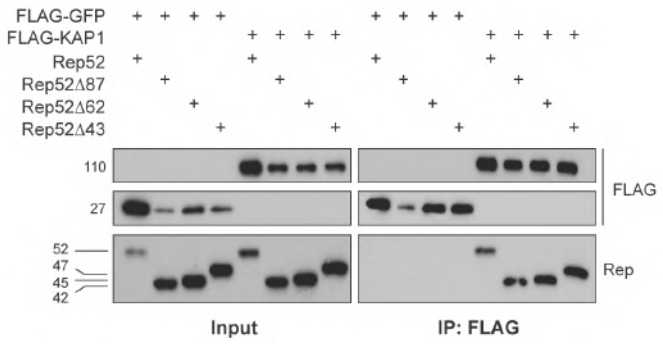
Extended Data Figure 3. Knockdown of CHD3 and SETDB1 leads to enhanced AAV2 replication and protein expression

a-c, AAV2 replication and protein expression in 293T cells depleted for CHD3 (siCHD3) and/or SETDB1 (siSETDB1). **a**, VP protein expression and depletion of CHD3 and SETDB1 analyzed by western blotting. **b**, Quantification of VP3 levels using ImageJ software. **c**, Viral genome replication analyzed by real time qPCR. Values are reported as mean±SEM, n=4.



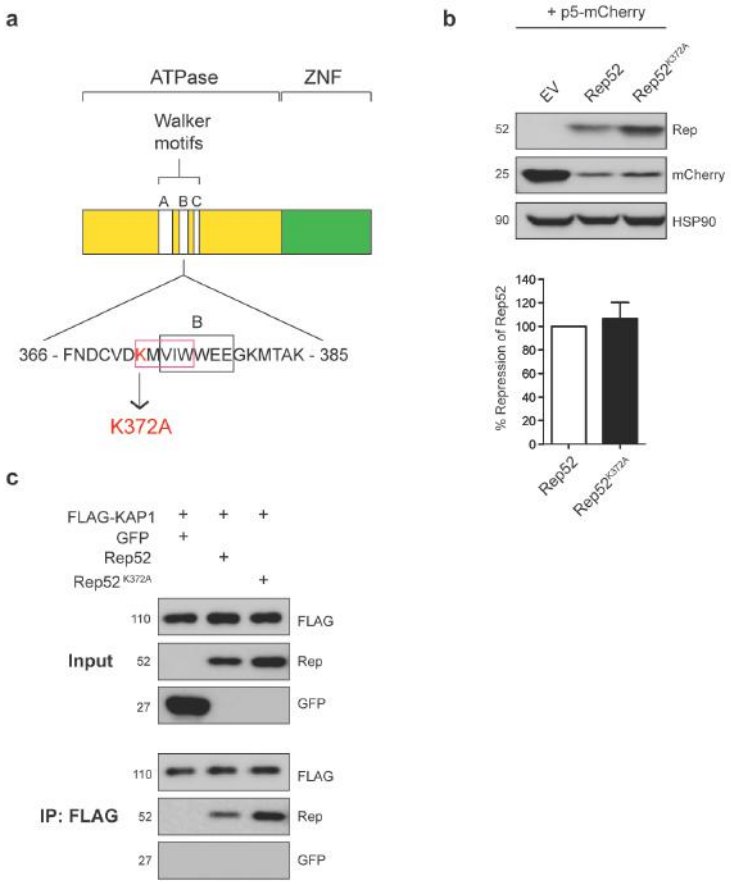
Extended Data Figure 4. Phosphorylation of KAP1-S824 by the Rep proteins is independent from ATM activation

a, p-KAP1-S824 localization in 293T cells pretreated with ATMi and expressing Rep78-GFP with and without Ad5 infection (left and right panel, respectively). **b-c,** p-KAP1-S824 levels analyzed in 293T cells pretreated with ATMi. **b**, cells were transfected with EV, Rep52, or Rep78. p-Chk2 was monitored to assess efficiency of ATM inhibition. p-KAP-S824 levels were normalized to Rep levels to correct for differences in transfection efficiency as a result of pretreatment with ATMi. **c**, Cells were infected with AAV2 and Ad5 as in Extended Data Figure 2. Values are reported as mean±SEM, n=4.



Extended Data Figure 5. Truncation of Rep C-terminal zinc finger domain does not affect Rep-KAP1 physical interaction.

Cross-linked co-IP for FLAG-tagged proteins from 293T cells expressing FLAG-KAP1, or a FLAG-GFP control, with full length Rep52 or each of the Rep52 C-terminal ZNF truncation mutants.



580

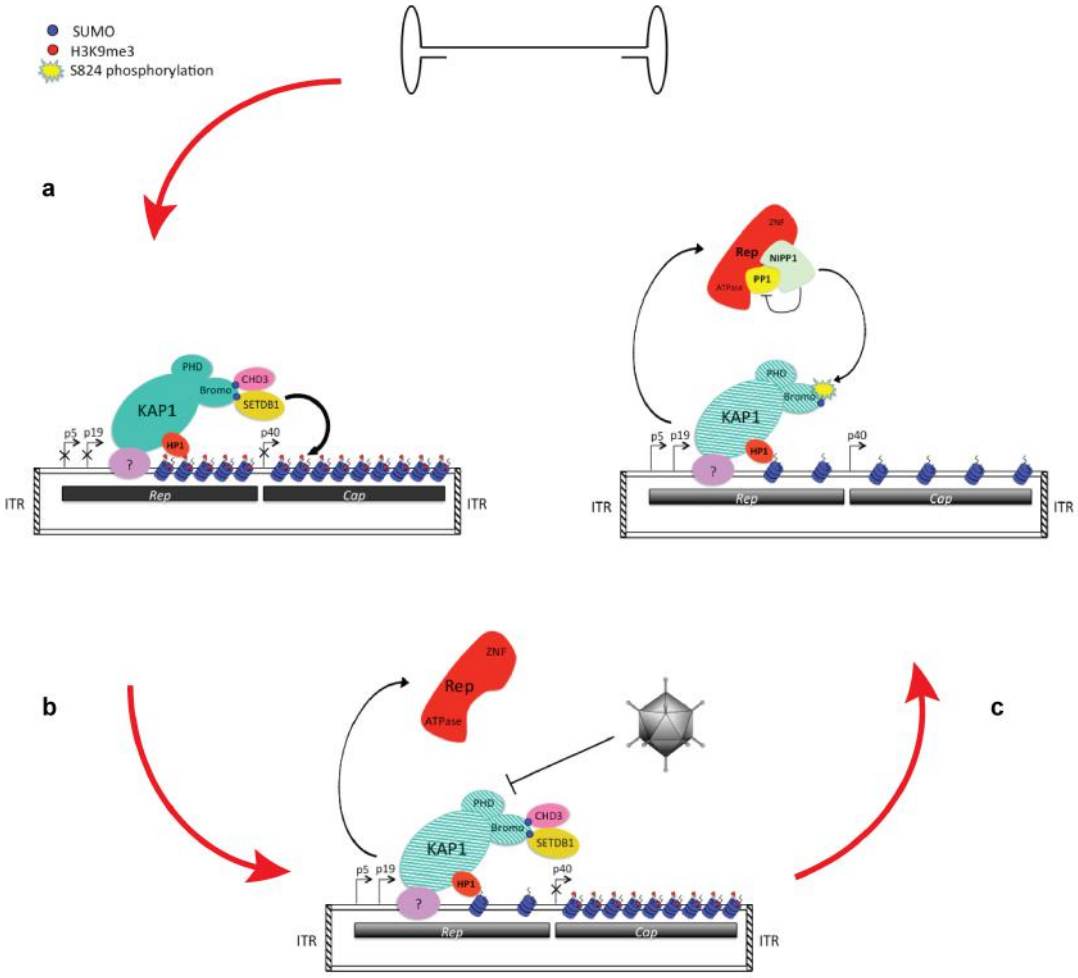
581

582

583 **Extended Data Figure 6. Validation of Rep-K372A PP1-binding mutant.**

584 **a**, Depiction of the PP1-binding site in the Rep ATPase domain. The Walker B motif is outlined
585 in black, and the partially overlapping consensus binding site is outlined in pink. **b**, Rep-mediated
586 repression of AAV2 p5 is dependent on a functional ATPase/helicase domain[28]. To verify
587 ATPase activity of Rep52^{K372A}, 293T cells expressing a p5-mCherry reporter construct with
588 Rep52 or Rep52^{K372A} were analyzed for p5 activity by western blotting for mCherry[18]. Protein
589 levels were quantified using Image J software. Values are reported as mean±SEM, n=3. **c**, Cross-
590 linked co-IP for FLAG-tagged proteins from 293T cells expressing FLAG-KAP1, or a FLAG-
591 GFP control, with the various Rep proteins.

592



593

594

595

596 **Extended Data Figure 7. Model for release of AAV2 from KAP1-mediated latency.**

597 **a**, Incoming AAV2 genomes undergo second-strand synthesis, concatamerization, and
598 chromatinization upon nuclear entry. KAP1 is recruited to the *rep* ORF via an unknown binding
599 partner where it forms a scaffold for the recruitment of SETDB1 and CHD3, leading to the
600 methylation of AAV2-associated histones. **b**, Partial degradation of KAP1 upon Ad5
601 superinfection triggers the initial upregulation of Rep. **c**, Binding of Rep to PP1 and NIPP1 leads
602 to enhanced levels of phosphorylated KAP1 and release of the repressive complex.